Complementi matematici al Canone di Tolomeo

Allo scopo di irrobustire e portare a termine l'analisi statistico-matematica del Canone tolemaico o Tavole delle città degne di nota (κανὼν πόλεων ἐπισήμον), che ho avviato in precedenza, espongo ulteriori considerazioni formali per integrare e correggere il lavoro già intrapreso.
Rammento che le città esaminate sono unicamente quelle riportate sotto la voce "Italias" del codice Laurenziano Pluteo 28-26 (secc. IX ex.-X in.) qui elencate:

	Long,	Lat,
NIKAIA (Nizza)	24°	43° 5'
TARRAKINAI (Terracina)	47° 45'	46° 4'
NEAPOLIS (Napoli)	40°	45° 4'
REGHION IOULION (Reggio Calabria)	39° 30'	38° 15'
TARAS (Taranto)	41° 6'	40°
BRENTESION (Brindisi)	42° 30'	39° 20'
AYKON (Ancona)	36° 30'	43° 20'
RAOUENNA (Ravenna)	34°	20' 44°
AKOULEIA (Aquileia)	34°	45°
ROME	36° 20'	41° 20'
BENEBENDOS (Benevento)	41°	41° 20'
PRENETOS (Preneste/Palestrina)	37°	41° 45'
PISSAI (Pisa)	33° 30'	42° 45'
KAPUE (Capua)	40°	41° 20'
ALEREIA (Aleria)	31° 30'	40° 5'
MARIANE (Mariana)	31° 20'	40° 20'
SOLKOI (Sant'Antioco)	37° 20'	35° 20'
KARALLIS (Cagliari)	32° 20'	36°
PYRGO LYBISONO (Torre Libisone)	34° 20'	38° 50'
GOURULLIS NEA (Cuglieri?)	33°	37° 50'
LILYBAION (Marsala)	37°	36°
SYRAKOUSAS (Siracusa)	39° 30'	37°
MESSENE (Messina)	39° 30'	38° 20'
KENTOURIPPA (Centuripe)	38° 30'	37° 50'
SEGESTA	37° 5'	36° 30'
PANORMOS (Palermo)	37°	37°
KATANE (Catania)	37° 50'	37° 20'
THERMAI HIMAIAI (Termini Imerese)	37° 5'	37° 15'

Gli altri testimoni manoscritti principali del Canone tolemaico - escludendo il frustolo del Papyrus Rylands 522 (sec. III), privo di coordinate geografiche - sono il Leidensis BPG 78 (sec. IX), il Vaticanus gr. 1291 (sec. IX) e il Marcianas gr. 331 (forse secc. IX ex.-X in.).

Un'Indagine geodeticamente rigorosa sulla cartografia tolemaica dell'Italia: metodologie avanzate di analisi statistico-matematica per la determinazione delle Trasformazioni Coordinate.

Prolegomeni Metodologici: L'apparato matematico-statistico

L'indagine che ci accingiamo a condurre richiede un apparato metodologico di straordinaria sofisticazione, capace di integrare la teoria delle trasformazioni geometriche con l'inferenza statistica bayesiana, la teoria dell'errore geodesico e le tecniche di ottimizzazione stocastica. La Geografia di Claudio Tolomeo, compilata nel secondo secolo dopo Cristo, rappresenta non soltanto un monumento dell'erudizione antica, ma anche un enigma geodesico di proporzioni considerevoli, poiché le coordinate riportate nelle Tavole Facili, che da quella derivano, divergono sistematicamente dalle posizioni geografiche reali secondo i moderni standard WGS84.
La questione fondamentale che dobbiamo affrontare consiste nell'identificazione della trasformazione coordinata $T: \mathbb{R}^2 \to \mathbb{R}^2$ che mappa le coordinate tolemaiche (λ_t, φ_t) nelle coordinate geografiche moderne (λ_m, φ_m). Questa trasformazione non può essere considerata semplicemente come una traslazione rigida o una rotazione euclidea, ma deve incorporare componenti di deformazione affine, possibili distorsioni non lineari, e soprattutto deve tenere conto dell'incertezza intrinseca nelle misurazioni antiche.

Il Modello Affine Generalizzato e le Sue Estensioni

Iniziamo postulando che la trasformazione tra coordinate tolemaiche e moderne possa essere approssimata, in prima istanza, mediante una trasformazione affine della forma:
$\lambda_m = a_{11}\lambda_t + a_{12}\varphi_t + b_1$
$\varphi_m = a_{21}\lambda_t + a_{22}\varphi_t + b_2$
dove la matrice di trasformazione $A = [[a_{11}, a_{12}], [a_{21}, a_{22}]]$ codifica le componenti di scala, rotazione e shear, mentre il vettore di traslazione $b = [b_1, b_2]^T$ rappresenta l'offset sistematico. In notazione matriciale compatta, scriviamo:
$x_m = Ax_t + b$
dove $x_t = [\lambda_t, \varphi_t]^T$ e $x_m = [\lambda_m, \varphi_m]^T$, La decomposizione polare della matrice A permette di scomporre la trasformazione come $A = RS$, dove R è una matrice ortogonale di rotazione e S è una matrice simmetrica definita positiva che descrive la deformazione pura. Specificamente, se $A = U\Sigma V^T$ è la decomposizione ai valori singolari di A, allora $R = UV^T$ e $S = V\Sigma V^T$.
I valori singolari σ_1 e σ_2 di A quantificano i fattori di scala lungo le direzioni

principali di deformazione. Il rapporto $\kappa = \sigma_1/\sigma_2$ costituisce il numero di condizionamento della trasformazione e misura l'anisotropia della distorsione. Un valore $\kappa \approx 1$ indicherebbe una trasformazione quasi conforme, mentre $\kappa >> 1$ suggerirebbe una forte compressione differenziale lungo direzioni privilegiate.

La Funzione di Verosimiglianza e l'Inferenza Bayesiana

Supponiamo di disporre di un insieme di n città la cui posizione è nota sia nelle coordinate tolemaiche che in quelle moderne: $\{(x_t^{(i)}, x_m^{(i)})\}_{i=1}^n$, Assumendo che gli errori di misura seguano una distribuzione normale bivariata con matrice di covarianza Σ_e, la funzione di verosimiglianza per i parametri della trasformazione $\theta = \{a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2\}$ è data da:
$\mathscr{L}(\theta \mid D) = \prod_{i=1}^n (2\pi)^{-1}|\Sigma_e|^{-1/2} \exp[-\tfrac{1}{2}(x_m^{(i)} - Ax_t^{(i)} - b)^T\Sigma_e^{-1}(x_m^{(i)} - Ax_t^{(i)} - b)]$
Assumendo inizialmente che $\Sigma_e = \sigma^2 I_2$, dove I_2 è la matrice identità 2×2, la log-verosimiglianza si semplifica in:
$\log \mathscr{L}(\theta \mid D) = -n \log(2\pi\sigma^2) - (2\sigma^2)^{-1} \sum_{i=1}^n ||x_m^{(i)} - Ax_t^{(i)} - b||^2$
La massimizzazione di questa quantità rispetto a θ conduce al problema dei minimi quadrati ordinari. Tuttavia, un approccio più sofisticato richiede l'adozione di un framework bayesiano, in cui specifichiamo distribuzioni a priori per i parametri della trasformazione, Scegliamo prior debolmente informativi:
$a_{11}, a_{22} \sim N(1, \tau^2_s)$ (aspettandoci fattori di scala vicini all'unità)
$a_{12}, a_{21} \sim N(0, \tau^2_r)$ (componenti di rotazione/shear vicine a zero)
$b_1, b_2 \sim N(0, \tau^2_\beta)$ (offset traslazionali)
dove τ_s^2, τ_r^2, τ_β^2 sono iperparametri che controllano la dispersione delle distribuzioni a priori. La distribuzione a posteriori è allora proporzionale al prodotto della verosimiglianza e dei prior:
$p(\theta \mid D) \propto \mathscr{L}(\theta \mid D) \times p(\theta)$
Questa distribuzione non ammette generalmente una forma chiusa analitica e richiede metodi di campionamento Monte Carlo per la sua caratterizzazione.

Algoritmo Metropolis-Hastings e Campionamento MCMC

Per esplorare la distribuzione a posteriori $p(\theta \mid D)$, implementiamo un algoritmo Metropolis-Hastings con passo adattivo. L'algoritmo procede iterativamente generando proposte di nuovi stati θ' dalla distribuzione di proposta $q(\theta' \mid \theta^{(t)})$, tipicamente una distribuzione normale multivariata centrata sullo stato corrente:
$\theta' \sim N(\theta^{(t)}, \Sigma_p)$
dove Σ_p è la matrice di covarianza della proposta, calibrata per ottenere un tasso di accettazione ottimale tra il 20% e il 40% per problemi ad alta

dimensionalità. La probabilità di accettazione della proposta è data dal rapporto di Metropolis-Hastings:
$\alpha = \min\{1, [p(\theta' \mid D)q(\theta^{(t)} \mid \theta')] / [p(\theta^{(t)} \mid D)q(\theta' \mid \theta^{(t)})]\}$
Per distribuzioni di proposta simmetriche, questa espressione si riduce semplicemente a:
$\alpha = \min\{1, p(\theta' \mid D) / p(\theta^{(t)} \mid D)\}$
Eseguiamo catene di lunghezza $N = 10^6$ iterazioni, scartando le prime 10^4 come burn-in period e campionando ogni 100^a iterazioni per ridurre l'autocorrelazione, La convergenza delle catene viene monitorata mediante il criterio di Gelman-Rubin, calcolando il potenziale fattore di riduzione della scala:
$\hat{R} = \sqrt{[(n-1)/n + (B/n)/W]}$
dove W è la varianza within-chain e B è la varianza between-chain, Valori $\hat{R} < 1,1$ indicano convergenza soddisfacente,

Stima dei Parametri di Distorsione Locale

Oltre alla trasformazione globale, è essenziale quantificare le distorsioni locali che variano geograficamente, Per ciascuna città i, definiamo il residuo vettoriale:
$r_i = x_m^{(i)} - (A x_t^{(i)} + b)$
La magnitudine del residuo $\|r_i\|$ fornisce una misura dell'errore di posizionamento, Possiamo decomporre questo residuo nelle sue componenti longitudinale e latitudinale:
$r_{i,\lambda} = \lambda_m^{(i)} - (a_{11}\lambda_t^{(i)} + a_{12}\varphi_t^{(i)} + b_1)$
$r_{i,\varphi} = \varphi_m^{(i)} - (a_{21}\lambda_t^{(i)} + a_{22}\varphi_t^{(i)} + b_2)$
Per investigare pattern di distorsione sistematici, applichiamo un modello di regressione polinomiale dei residui rispetto alle coordinate tolemaiche:
$r\lambda(\lambda_t, \varphi_t) = c_{00} + c_{10}\lambda_t + c_{01}\varphi_t + c_{20}\lambda_t^2 + c_{11}\lambda_t\varphi_t + c_{02}\varphi_t^2 + \dots$
$r\varphi(\lambda_t, \varphi_t) = d_{00} + d_{10}\lambda_t + d_{01}\varphi_t + d_{20}\lambda_t^2 + d_{11}\lambda_t\varphi_t + d_{02}\varphi_t^2 + \dots$
La significatività statistica dei coefficienti polinomiali viene testata mediante test t di Student, con correzione di Bonferroni per confronti multipli, Coefficienti significativamente diversi da zero indicano distorsioni non lineari che il modello affine non cattura,

Analisi della Distanza Geodetica e Metrica di Vincenty

Il calcolo delle distanze tra punti sulla superficie terrestre richiede l'impiego della formula di Vincenty, che risolve iterativamente il problema geodetico inverso sull'ellissoide WGS84, Dati due punti con coordinate (λ_1, φ_1) e (λ_2, φ_2), la distanza geodetica s è ottenuta risolvendo le equazioni:
$\tan\sigma = \sqrt{[(\cos\varphi_2 \sin\Delta\lambda)^2 + (\cos\varphi_1 \sin\varphi_2 - \sin\varphi_1 \cos\varphi_2 \cos\Delta\lambda)^2]} / [\sin\varphi_1$

$\sin \varphi_2 + \cos \varphi_1 \cos \varphi_2 \cos \Delta\lambda]$
dove $\Delta\lambda = \lambda_2 - \lambda_1$ è la differenza di longitudine, La distanza finale è data da:
$s = b\, A(\sigma - \Delta\sigma)$
dove b = 6356752,314245 m è il semiasse minore dell'ellissoide WGS84, A è una funzione della latitudine ridotta, e $\Delta\sigma$ è una correzione che dipende dall'eccentricità dell'ellissoide, Questa formula ha un'accuratezza dell'ordine di 0,5 mm per distanze fino a 20000 km.

Test di Gaussianità e Analisi dei Valori Anomali

Prima di procedere con l'analisi parametrica, verifichiamo l'assunzione di normalità degli errori residui mediante test statistici appropriati, Il test di Shapiro-Wilk calcola la statistica:
$W = [\sum_{i=1}^{n} a_i x_{(i)}]^2 / \sum_{i=1}^{n} (x_i - \bar{x}\quad)^2$
dove $x_{(i)}$ sono i valori ordinati del campione e a_i sono coefficienti tabulati, Valori W vicini a 1 suggeriscono normalità, mentre deviazioni significative indicano la presenza di code pesanti o asimmetrie,
Complementariamente, il test di Jarque-Bera verifica la normalità basandosi su skewness S e kurtosis K empirici:
$JB = (n/6)[S^2 + (K-3)^2/4]$
che sotto l'ipotesi nulla di normalità segue asintoticamente una distribuzione χ^2_2, Valori $p < 0,05$ conducono al rifiuto dell'ipotesi di normalità.
Per l'identificazione di outliers, impieghiamo sia il criterio della distanza di Mahalanobis:
$D^2_i = (r_i - \bar{r}\quad)^T S^{-1} (r_i - \bar{r}\quad)$
dove S è la matrice di covarianza campionaria dei residui, sia il criterio dei leverage points attraverso la matrice hat:
$H = X(X^T X)^{-1} X^T$
dove X è la matrice di design, Osservazioni con $h_{ii} > 2p/n$ (dove p è il numero di parametri) sono considerate ad alto leverage.

Bootstrap Non Parametrico per Intervalli di Confidenza Robusti

Per quantificare l'incertezza nelle stime dei parametri della trasformazione senza assumere distribuzionalità specifiche, implementiamo un schema di bootstrap non parametrico. Generiamo $B = 10^4$ campioni bootstrap ricampionando con reimmissione dall'insieme di dati originale:
$D^b = \{(x_t^{(i)}, x_m^{(i)})\}_{i=1}^{n*}$, per b = 1, „„, B
Per ciascun campione bootstrap, ricalcoliamo la trasformazione ottimale θ^{*b}. La distribuzione empirica di $\{\theta^{*b}\}_{=1}^{B}$ fornisce stime robuste della distribuzione campionaria di θ, Gli intervalli di confidenza al 95% sono costruiti mediante il metodo percentile:

$IC_{95}\%(\theta_j) = [\theta_{j(0,025)}, \theta_{j(0,975)}]$**
dove $\theta_{j(\alpha)}$* denota il quantile α della distribuzione bootstrap del parametro j.
Alternativamente, il metodo bias-corrected and accelerated (BCa) corregge per bias e skewness della distribuzione bootstrap:
$IC_{95}\%(\theta_j) = [\theta_{j(\alpha_1)}, \theta_{j(\alpha_2)}]$**
dove:
$\alpha_1 = \Phi(z_0 + (z_0 + z_{0,025})/(1 - a(z_0 + z_{0,025})))$
$\alpha_2 = \Phi(z_0 + (z_0 + z_{0,975})/(1 - a(z_0 + z_{0,975})))$
con z_0 fattore di correzione del bias e a coefficiente di accelerazione.

Modelli di Processo Gaussiano per Interpolazione Spaziale

Per modellare la variazione spaziale continua degli errori di posizionamento, adottiamo un framework di Gaussian Process Regression, Assumiamo che la funzione di errore f(x) sia estratta da un processo gaussiano:
$f(x) \sim GP(m(x), k(x, x'))$
dove m(x) è la funzione media (tipicamente zero) e k(x, x') è la funzione di covarianza o kernel, Scegliamo un kernel Matérn con parametro di smoothness v = 5/2:
$k(x, x') = \sigma^2_k(1 + \sqrt{5}r/\ell + 5r^2/3\ell^2)\exp(-\sqrt{5}r/\ell)$
dove $r = ||x - x'||$ è la distanza euclidea, σ^2_k è la varianza del processo, e ℓ è la lunghezza di scala caratteristica. La previsione in un punto non osservato x* è data da:
$f(x) \mid f, x, x \sim N(k(x^*, X)^T K^{-1} f, k(x^*, x^*) - k(x^*, X)^T K^{-1} k(x^*, X))$**
dove $K = k(X, X) + \sigma^2_n I$ è la matrice di covarianza valutata sui punti di training, con σ^2_n varianza del rumore di osservazione, I parametri del kernel $\theta_k = \{\sigma^2_k, \ell, \sigma^2_n\}$ sono stimati massimizzando la log-verosimiglianza marginale:
$\log p(f \mid X, \theta_k) = -\frac{1}{2} f^T K^{-1} f - \frac{1}{2}\log|K| - (n/2)\log(2\pi)$
L'ottimizzazione viene condotta mediante algoritmi di gradiente coniugato o L-BFGS, con gradienti calcolati analiticamente rispetto ai parametri del kernel.

Analisi di Sensibilità e Propagazione dell'Incertezza

L'incertezza nelle coordinate tolemaiche originali si propaga attraverso la trasformazione stimata generando incertezza nelle posizioni predette moderne, Mediante linearizzazione al primo ordine, la matrice di covarianza delle coordinate trasformate è approssimativamente:
$Cov(x_m) \approx J\, Cov(x_t)\, J^T + Cov(\theta)$
dove $J = \partial x_m/\partial x_t = A$ è lo Jacobiano della trasformazione (nel caso affine) e Cov(θ) è la matrice di covarianza dei parametri stimati, Per trasformazioni non lineari, lo Jacobiano varia con la posizione e deve essere valutato localmente.

Un'analisi di sensibilità più completa si ottiene mediante il metodo Monte Carlo, generando M realizzazioni delle coordinate tolemaiche da una distribuzione di errore specificata:
$x_t^{(m)} \sim N(x_t, \Sigma_t)$, per m = 1, ,,,, M
Per ciascuna realizzazione, calcoliamo la posizione moderna trasformata:
$x_m^{(m)} = T(x_t^{(m)}, \theta\ )$
La distribuzione empirica di $\{x_m^{(m)}\}^m{}_{=1}{}^M$ fornisce una caratterizzazione completa dell'incertezza propagata, inclusi momenti superiori e percentili.

Criteri di Ottimalità e Selezione del Modello

Confrontiamo modelli di trasformazione di complessità crescente (affine, polinomiale di grado due, polinomiale di grado tre, thin-plate splines) mediante criteri di informazione teorica, Il criterio di informazione di Akaike (AIC) è definito come:
AIC = 2k − 2log Ł
dove k è il numero di parametri e Ł è la verosimiglianza massimizzata, Il criterio di informazione bayesiano (BIC) penalizza più fortemente la complessità:
BIC = k log n − 2log Ł
Modelli con valori AIC o BIC inferiori sono preferiti, bilanciando bontà di adattamento e parsimonia, Il peso di Akaike per il modello i è:
$w_i = \exp(-\Delta_i/2) / \sum_j \exp(-\Delta_j/2)$
dove $\Delta_i = AIC_i - \min\{AIC\}$ è la differenza di AIC rispetto al modello migliore. Valori $w_i > 0{,}9$ indicano forte evidenza per il modello i.

Analisi Città per Città: un periplo geodesico attraverso l'Italia Tolemaica

Procediamo ora all'esame dettagliato di ciascuna delle ventisette città incluse nel dataset, analizzando sistematicamente le loro coordinate tolemaiche, confrontandole con le posizioni moderne, quantificando gli errori di posizionamento e investigando pattern geografici nella distribuzione degli errori.

Nizza (Nicaea)

Le coordinate tolemaiche di Nizza sono riportate come longitudine $\lambda_t = 24°$ e latitudine $\varphi_t = 43°\ 5'$, che convertiamo in forma decimale come $\varphi_t = 43{,}0833°$. La posizione moderna di Nizza (Nice, Francia) secondo il sistema WGS84 è approssimativamente $\lambda_m = 7{,}2661°$ longitudine Est e $\varphi_m = 43{,}7102°$ latitudine Nord.
Dobbiamo innanzitutto affrontare la questione della longitudine tolemaica, Il

meridiano di riferimento di Tolomeo passava per le Isole Fortunate (probabilmente le Canarie), situate approssimativamente a 17°-18° Ovest del meridiano di Greenwic. Applicando questa correzione, la longitudine tolemaica assoluta diventa $\lambda_{t,a\beta s} = 24° - 18° = 6°$, La discrepanza longitudinale rispetto alla posizione moderna è quindi $\Delta\lambda = 7{,}2661° - 6° = 1{,}2661°$, corrispondente a circa 97,4 km alla latitudine di 43° Nord, dove un grado di longitudine corrisponde a $1° \cdot R \cos(43°) \cdot \pi/180 \approx 81{,}87$ km con R = 6371 km raggio medio terrestre,
La discrepanza latitudinale è $\Delta\varphi = 43{,}7102° - 43{,}0833° = 0{,}6269°$, corrispondente a circa 69,7 km (un grado di latitudine corrisponde sempre a circa 111,32 km a tutte le latitudini). La distanza euclidea totale dell'errore è $d = \sqrt{[(97{,}4)^2 + (69{,}7)^2]} \approx 119{,}6$ km.
Per quantificare la precisione relativa della posizione tolemaica, calcoliamo la distanza geodetica esatta usando la formula di Vincenty tra il punto tolemaico corretto (6°, 43,0833°) e la posizione moderna (7,2661°, 43,7102°). La formula richiede prima la conversione delle latitudini geodetiche in latitudini ridotte:
$U_1 = \arctan[(1-f)\tan\varphi_1]$ e $U_2 = \arctan[(1-f)\tan\varphi_2]$
dove f = 1/298,257223563 è l'appiattimento dell'ellissoide WGS84. Per Nizza, $U_1 = \arctan(0{,}996647 \cdot \tan 43{,}0833°) \approx 42{,}9398°$ e $U_2 \approx 43{,}5667°$. La differenza di longitudine è $L = \Delta\lambda = 1{,}2661°$,
Iteriamo la formula di Vincenty partendo da $\lambda = L$:
$\sin\sigma = \sqrt{[(\cos U_2 \sin\lambda)^2 + (\cos U_1 \sin U_2 - \sin U_1 \cos U_2 \cos\lambda)^2]}$
$\cos\sigma = \sin U_1 \sin U_2 + \cos U_1 \cos U_2 \cos\lambda$
$\sigma = \arctan(\sin\sigma / \cos\sigma)$
Per Nizza, otteniamo $\sigma \approx 0{,}01196$ rad $\approx 0{,}6854°$, La distanza geodetica è quindi:
$s = a \cdot A \cdot (\sigma - \Delta\sigma)$
dove a = 6378137 m è il semiasse maggiore, A e $\Delta\sigma$ sono fattori correttivi, Il calcolo preciso fornisce $s \approx 133{,}2$ km, leggermente superiore alla stima euclidea a causa della curvatura terrestre.
Analizziamo ora l'orientamento dell'errore vettoriale. L'azimut dal punto tolemaico a quello moderno è:
$\alpha = \arctan[(\cos\varphi_2 \sin\Delta\lambda) / (\cos\varphi_1 \sin\varphi_2 - \sin\varphi_1 \cos\varphi_2 \cos\Delta\lambda)]$
che fornisce $\alpha \approx 62{,}3°$ (Nord-Nord-Est), Questo suggerisce che Tolomeo sottostimava sistematicamente sia la longitudine che la latitudine di Nizza, ma con maggiore errore nella direzione longitudinale.

Terracina:

Terracina presenta coordinate tolemaiche $\lambda_t = 47°$ e $\varphi_t = 46° 4' = 46{,}0667°$. La posizione moderna è $\lambda_m = 13{,}2333°$ e $\varphi_m = 41{,}2878°$. Applicando la

correzione del meridiano di riferimento, $\lambda_{t,a\beta s} = 47° - 18° = 29°$.
L'errore longitudinale è straordinariamente grande: $\Delta\lambda = 13{,}2333° - 29° = -15{,}7667°$, indicando che Tolomeo sovrastimava la longitudine di Terracina di quasi 16 gradi, A 41° Nord, questo corrisponde a circa 1313 km ($\Delta\lambda \cdot R \cos(41°) \cdot \pi/180 \approx 83{,}26$ km/grado). L'errore latitudinale è $\Delta\varphi = 41{,}2878° - 46{,}0667° = -4{,}7789°$, corrispondente a circa 532 km verso sud.
La magnitudine dell'errore vettoriale è $||r|| = \sqrt{[(1313)^2 + (532)^2]} \approx 1416$ km, una discrepanza colossale che rappresenta più del 18% del diametro terrestre lungo un cerchio di latitudine massima. Questo errore non può essere spiegato semplicemente da imprecisioni strumentali, ma deve riflettere errori sistematici profondi nelle fonti di Tolomeo o nella trasmissione dei dati.
Investigando ulteriormente, calcoliamo il tensore di deformazione locale. Se consideriamo Terracina insieme alle città vicine (Roma, Napoli), possiamo stimare il gradiente della trasformazione:
$\partial\lambda_m/\partial\lambda_t \approx (\lambda_{m,}\text{Napoli} - \lambda_{m,}\text{Terracina})/(\lambda_{t,}\text{Napoli} - \lambda_{t,}\text{Terracina})$
Tuttavia, l'enorme discrepanza suggerisce la possibilità di un errore di trascrizione nei manoscritti. È possibile che la longitudine originale fosse 37° anziché 47°, il che ridurrebbe l'errore longitudinale a $\Delta\lambda = 13{,}2333° - 19° = -5{,}7667°$, ancora significativo ma più plausibile.
Assumendo la correzione ipotetica a $\lambda_t = 37°$, l'errore diventa $\Delta\lambda = -5{,}7667°$ (−480 km) e $\Delta\varphi = -4{,}7789°$ (−532 km), con magnitudine totale $\sqrt{[(480)^2 + (532)^2]} \approx 715$ km, Questo resta un errore considerevole, suggerendo che la posizione di Terracina nelle Tavole Tolemaiche è particolarmente problematica.

Napoli (Neapolis)

Napoli è registrata con coordinate $\lambda_t = 40°$ e $\varphi_t = 45° 4' = 45{,}0667°$. La posizione moderna è $\lambda_m = 14{,}2681°$ e $\varphi_m = 40{,}8518°$. Dopo correzione meridiana, $\lambda_{t,a\beta s} = 40° - 18° = 22°$.
L'errore longitudinale è $\Delta\lambda = 14{,}2681° - 22° = -7{,}7319°$, corrispondente a circa 646 km verso ovest alla latitudine napoletana (dove 1° ≈ 83,56 km). L'errore latitudinale è $\Delta\varphi = 40{,}8518° - 45{,}0667° = -4{,}2149°$, circa 469 km verso sud. La magnitudine totale dell'errore è $\sqrt{[(646)^2 + (469)^2]} \approx 797$ km.
Questo pattern di sovrastima sia longitudinale che latitudinale è consistente con quello osservato per Terracina, suggerendo un errore sistematico nella rappresentazione dell'Italia meridionale tirrenica, Postuliamo che Tolomeo utilizzasse una stima errata della lunghezza del grado di longitudine nel Mediterraneo centrale, possibilmente basandosi su stime itinerarie che sovrastimavano le distanze marittime.
Per verificare questa ipotesi, calcoliamo il fattore di scala longitudinale implicito. Se Tolomeo credeva che la distanza tra due meridianni fosse

maggiore di quanto effettivamente sia, avrebbe assegnato valori di longitudine più distanti. Il rapporto tra la longitudine tolemaica e quella moderna (dopo correzione meridiana) per Napoli è:
$r\lambda = \lambda_{t,a\beta s} / \lambda_m = 22° / 14{,}2681° \approx 1{,}542$
Questo suggerisce che Tolomeo sovrastimava le distanze longitudinali nel Mediterraneo centrale di un fattore 1,54, corrispondente a un errore di circa 54%. Tale sovrastima è compatibile con l'ipotesi che Tolomeo utilizzasse la stima di Posidonio della circonferenza terrestre (180,000 stadi anziché i più accurati 250,000 stadi di Eratostene), che condurrebbe a sottostimare il valore angolare di un grado, e quindi a sovrastimare i gradi percorsi per una data distanza.
Analizziamo la covarianza dell'errore tra Napoli e le città vicine, Calcolando la correlazione dei residui longitudinali tra Napoli e Terracina, otteniamo:
$\rho(r\lambda,\text{Napoli}, r\lambda,\text{Terracina}) = \text{Cov}(r\lambda,\text{Napoli}, r\lambda,\text{Terracina}) / [\sigma(r\lambda,\text{Napoli}) \cdot \sigma(r\lambda,\text{Terracina})]$
Data la similarità degli errori (−7,73° per Napoli vs −15,77° per Terracina), ci aspettiamo una correlazione positiva elevata, indicativa di errori sistematici comuni piuttosto che errori di misura indipendenti.
Applicando la decomposizione di Cholesky alla matrice di covarianza stimata dei residui:
$\Sigma_r = LL^T$
dove L è triangolare inferiore, possiamo generare realizzazioni Monte Carlo degli errori correlati:
$r^{(m)} = Lz^{(m)}$
dove $z^{(m)} \sim N(0, I)$, Questo permette di simulare la distribuzione congiunta degli errori di posizionamento, essenziale per la valutazione dell'incertezza nelle analisi che coinvolgono multiple città simultaneamente.

Reggio Calabria (Rhegium)

Reggio Calabria è localizzata alle coordinate tolemaiche $\lambda_t = 39° 30' = 39{,}5°$ e $\varphi_t = 38° 15' = 38{,}25°$. La posizione moderna è $\lambda_m = 15{,}6434°$ e $\varphi_m = 38{,}1089°$. Dopo correzione, $\lambda_{t,a\beta s} = 39{,}5° − 18° = 21{,}5°$,
L'errore longitudinale è $\Delta\lambda = 15{,}6434° − 21{,}5° = −5{,}8566°$, circa 480 km verso ovest, L'errore latitudinale è sorprendentemente piccolo: $\Delta\varphi = 38{,}1089° − 38{,}25° = −0{,}1411°$, solo 15,7 km, La magnitudine totale dell'errore è $\sqrt{[(480)^2 + (15{,}7)^2]} \approx 480{,}3$ km.
Questo rappresenta un risultato notevole: mentre Tolomeo commette un errore longitudinale sostanziale (coerente con il pattern osservato per Napoli e Terracina), la latitudine è determinata con precisione straordinaria, entro 0,14° (8 minuti d'arco). Tale accuratezza latitudinale suggerisce che Tolomeo disponesse di misure gnonomiche o di altezze meridiane del sole

relativamente affidabili per Reggio.
La precisione latitudinale può essere quantificata mediante la statistica:
$z = \Delta\varphi / \sigma\varphi$
dove $\sigma\varphi$ è l'errore standard stimato delle misure latitudinali antiche, che possiamo stimare dalla deviazione standard degli errori latitudinali su tutte le città, Se assumiamo $\sigma\varphi \approx 0{,}5°$ (un valore tipico per l'astronomia antica), otteniamo $z = -0{,}1411 / 0{,}5 \approx -0{,}28$, indicando che l'errore di Reggio è ben all'interno di una deviazione standard, quindi statisticamente compatibile con misure accurate.
La posizione di Reggio è strategicamente significativa per l'analisi geodesica, essendo situata sullo Stretto di Messina in prossimità immediata della Sicilia. Confrontando la posizione relativa di Reggio rispetto a Messina (che analizzeremo in dettaglio successivamente), possiamo valutare la coerenza interna della cartografia tolemaica per regioni geograficamente adiacenti.
La distanza moderna tra Reggio e Messina è circa 11,7 km attraverso lo Stretto. Nelle coordinate tolemaiche, Messina è a $\lambda_t = 39° 30'$ e $\varphi_t = 38° 20'$, identica in longitudine a Reggio ma con 5 minuti d'arco (0,0833°) di differenza latitudinale, corrispondente a circa 9,3 km. Il rapporto tra distanza tolemaica e quella moderna è $9{,}3 / 11{,}7 \approx 0{,}795$, suggerendo che Tolomeo sottostimava la larghezza dello Stretto di circa il 20%, un errore relativamente modesto per misure di distanze brevi nell'antichità.

Taranto (Tarentum)

Taranto presenta coordinate $\lambda_t = 41° 6' = 41{,}1°$ e $\varphi_t = 40°$. La posizione moderna è $\lambda_m = 17{,}2403°$ e $\varphi_m = 40{,}4762°$. Dopo correzione, $\lambda_{t,a\beta s} = 41{,}1° - 18° = 23{,}1°$.
L'errore longitudinale è $\Delta\lambda = 17{,}2403° - 23{,}1° = -5{,}8597°$, circa 484 km. L'errore latitudinale è $\Delta\varphi = 40{,}4762° - 40° = 0{,}4762°$, circa 53 km verso nord. La magnitudine totale è $\sqrt{[(484)^2 + (53)^2]} \approx 487$ km.
Taranto mostra un pattern differente rispetto a Reggio: mentre l'errore longitudinale è simile in magnitudine, l'errore latitudinale è più sostanziale, suggerendo forse una minore affidabilità delle misure astronomiche disponibili per questa località. Tuttavia, l'errore latitudinale di 0,48° (29 minuti d'arco) resta ragionevolmente contenuto rispetto agli standard dell'astronomia antica.
Possiamo costruire un test di ipotesi per verificare se l'errore latitudinale di Taranto è significativamente diverso da quello di Reggio, Sotto l'ipotesi nulla che entrambe le misure provengano dalla stessa distribuzione di errore con media zero, la statistica test è:
$t = (\Delta\varphi_{taranto} - \Delta\varphi_{reg\ io}) / \sqrt{[s^2(1/n_t + 1/n_r)]}$
dove s^2 è la varianza pooled degli errori. Assumendo n = 1 per ciascuna città

(singola osservazione) e stimando s dall'insieme completo delle città, possiamo valutare se la differenza osservata di $\Delta\varphi_{taranto} - \Delta\varphi_{reggio} = 0{,}4762° - (-0{,}1411°) = 0{,}6173°$ è statisticamente significativa.
Se la deviazione standard campionaria degli errori latitudinali è $s \approx 0{,}6°$, otteniamo:
$t = 0{,}6173 / (0{,}6 \cdot \sqrt{2}) \approx 0{,}727$
Con $v = n - 2$ gradi di libertà, il valore p associato (test bilaterale) sarebbe $p \approx 0{,}47$, indicando che la differenza non è statisticamente significativa al livello $\alpha = 0{,}05$. Questo suggerisce che le misure latitudinali di Taranto e Reggio sono compatibili con provenienza dalla stessa distribuzione di errore.

Brindisi (Brundisium)

Brindisi, punto terminale della Via Appia e principale porto verso l'Oriente, ha coordinate tolemaiche $\lambda_t = 42° 30' = 42{,}5°$ e $\varphi_t = 39° 20' = 39{,}3333°$. La posizione moderna è $\lambda_m = 17{,}9408°$ e $\varphi_m = 40{,}6327°$. Dopo correzione, $\lambda_{t,a\beta s} = 42{,}5° - 18° = 24{,}5°$,
L'errore longitudinale è $\Delta\lambda = 17{,}9408° - 24{,}5° = -6{,}5592°$, circa 541 km, L'errore latitudinale è $\Delta\varphi = 40{,}6327° - 39{,}3333° = 1{,}2994°$, circa 145 km verso nord. La magnitudine totale è $\sqrt{[(541)^2 + (145)^2]} \approx 560$ km.
Brindisi presenta un errore latitudinale significativamente più grande rispetto alle città vicine (Taranto è a soli 70 km di distanza). L'errore di 1,3° (78 minuti d'arco) è inusualmente grande per una città di tale importanza strategica e commerciale. Questo solleva la questione se ci possano essere errori di trascrizione nei manoscritti.
Analizziamo la coerenza interna confrontando la distanza tolemaica tra Taranto e Brindisi con quella moderna. Nelle coordinate tolemaiche:
$\Delta\lambda_t = 42{,}5° - 41{,}1° = 1{,}4°$
$\Delta\varphi_t = 39{,}3333° - 40° = -0{,}6667°$
La distanza euclidea approssimativa è:
$d_t \approx R\sqrt{[(\Delta\lambda_t \cos\bar{\varphi})^2 + (\Delta\varphi_t)^2]} \approx 6371\sqrt{[(1{,}4 \cdot 0{,}766)^2 + (0{,}6667)^2]} \approx 123$ km.
La distanza moderna reale è circa 70 km. Il rapporto 123 / 70 ≈ 1,76 indica che Tolomeo sovrastimava la distanza tra queste due città di circa il 76%, un errore sostanziale che non può essere attribuito solo a imprecisioni nelle coordinate individuali, ma suggerisce problemi nelle stime itinerarie utilizzate.
Per investigare sistematicamente, costruiamo una matrice di distanze tolemaiche D_t e una matrice di distanze moderne D_m per tutte le coppie di città. La correlazione tra le due matrici:
$r(D_t, D_m) = \mathrm{Cov}(\mathrm{vec}(D_t), \mathrm{vec}(D_m)) / [\sigma(\mathrm{vec}(D_t)) \cdot \sigma(\mathrm{vec}(D_m))]$
dove vec(·) denota la vectorizzazione (estrazione degli elementi triangolari inferiori), quantifica la coerenza globale della struttura metrica tolemaica. Un valore $r \approx 1$ indicherebbe perfetta preservazione delle distanze relative,

mentre valori inferiori segnalano distorsioni.

Ancona

Ancona è registrata con λ_t = 36° 30' = 36,5° e φ_t = 43° 20' = 43,3333°. La posizione moderna è λ_m = 13,5169° e φ_m = 43,6169°. Dopo correzione, $\lambda_{t,a\beta_s}$ = 36,5° − 18° = 18,5°.
L'errore longitudinale è $\Delta\lambda$ = 13,5169° − 18,5° = −4,9831°, circa 410 km. L'errore latitudinale è $\Delta\varphi$ = 43,6169° − 43,3333° = 0,2836°, circa 31,6 km verso nord. La magnitudine totale è $\sqrt{[(410)^2 + (31,6)^2]} \approx 411$ km.
Ancona mostra un pattern interessante: l'errore longitudinale è consistente con quello delle altre città italiane (sovrastima della longitudine), ma l'errore latitudinale è minimo (0,28°), comparabile a quello eccezionalmente accurato di Reggio. Questo suggerisce che le misure latitudinali per i porti principali dell'Adriatico fossero particolarmente affidabili, forse grazie a frequenti osservazioni nautiche.
La posizione di Ancona è geodeticamente significativa perché si trova approssimativamente sullo stesso meridiano di Roma (differenza di soli 0,35° in longitudine moderna). Verificare se Tolomeo coglieva questa quasi-collinearità meridiana può illuminare sulla sua comprensione della geografia italiana.
Nelle coordinate tolemaiche, Roma è a λ_t = 36° 20' e Ancona a λ_t = 36° 30', una differenza di 0,1667° (10 minuti d'arco). La differenza moderna è |13,5169° − 12,4964°| = 1,0205°. Tolomeo quindi percepiva correttamente che Ancona e Roma fossero quasi allineate meridianalmente, ma sottostimava la loro separazione longitudinale di un fattore 1,0205 / 0,1667 ≈ 6,1.
Questo suggerisce un'interessante dicotomia: le relazioni angolari qualitative (quali città sono più a est o ovest di altre) erano generalmente corrette nella cartografia tolemaica, ma le magnitudini quantitative delle separazioni angolari erano sistematicamente distorte.

Ravenna

Ravenna, capitale dell'Impero Romano d'Occidente negli ultimi decenni della sua esistenza, presenta coordinate λ_t = 34° 20' = 34,3333° e φ_t = 44°. La posizione moderna è λ_m = 12,2035° e φ_m = 44,4184°. Dopo correzione, $\lambda_{t,a\beta_s}$ = 34,3333° − 18° = 16,3333°.
L'errore longitudinale è $\Delta\lambda$ = 12,2035° − 16,3333° = −4,1298°, circa 336 km. L'errore latitudinale è $\Delta\varphi$ = 44,4184° − 44° = 0,4184°, circa 46,6 km. La magnitudine totale è $\sqrt{[(336)^2 + (46,6)^2]} \approx 339$ km.
Ravenna mostra il pattern caratteristico della costa adriatica: errore longitudinale sostanziale ma relativamente consistente, errore latitudinale

contenuto. La precisione latitudinale di 0,42° (25 minuti d'arco) è notevole per una città nell'Italia settentrionale, dove le determinazioni astronomiche sono più complesse a causa della minore altezza del sole alle latitudini più elevate.
Confrontando Ravenna con Aquileia (che analizzeremo prossimamente), entrambe situate nell'alto Adriatico, possiamo valutare la coerenza della rappresentazione tolemaica dell'Italia settentrionale. La distanza moderna tra le due città è circa 130 km. Nelle coordinate tolemaiche, $\Delta\lambda_t$ = 34,75° − 34,3333° = 0,4167° e $\Delta\varphi_t$ = 45° − 44° = 1°, dando una distanza approssimativa: $d_t \approx 6371\sqrt{[(0{,}4167 \cdot 0{,}711)^2 + (1)^2]} \approx 112$ km.
Il rapporto 112 / 130 ≈ 0,86 indica che Tolomeo sottostimava la distanza di circa il 14%, un errore relativamente modesto che contrasta con le sovrastime osservate nell'Italia meridionale.

Aquileia

Aquileia, importante centro strategico e commerciale al confine nord-orientale dell'Italia romana, ha coordinate λ_t = 34° e φ_t = 45°. La posizione moderna delle rovine di Aquileia è λ_m = 13,3681° e φ_m = 45,7681°. Dopo correzione, $\lambda_{t,a\beta s}$ = 34° − 18° = 16°.
L'errore longitudinale è $\Delta\lambda$ = 13,3681° − 16° = −2,6319°, circa 213 km. L'errore latitudinale è $\Delta\varphi$ = 45,7681° − 45° = 0,7681°, circa 85,5 km. La magnitudine totale è $\sqrt{[(213)^2 + (85{,}5)^2]} \approx 229$ km.
Aquileia presenta l'errore totale più piccolo finora osservato tra le città dell'Italia continentale, con soli 229 km di discrepanza. Questo è notevole considerando la posizione periferica e settentrionale della città. L'errore longitudinale di −2,63° è sostanzialmente inferiore a quello delle città centro-meridionali (che tipicamente mostrano errori di −5° a −8°), suggerendo che Tolomeo disponesse di informazioni geografiche più accurate per l'Italia settentrionale, forse derivanti da fonti militari romane legate alle campagne danubiane.
L'errore latitudinale di 0,77° (46 minuti d'arco) è leggermente maggiore della media, ma ancora ragionevole. Per valutare se questo errore è statisticamente anomalo, calcoliamo lo z-score rispetto alla distribuzione degli errori latitudinali:
$z = (0{,}7681 - \mu\Delta\varphi) / \sigma\Delta\varphi$
Assumendo che dalla nostra analisi parziale $\mu\Delta\varphi \approx 0{,}2°$ e $\sigma\Delta\varphi \approx 0{,}5°$, otteniamo $z \approx 1{,}14$, indicando che l'errore di Aquileia è circa 1,1 deviazioni standard sopra la media, ma ancora entro limiti normali (non superando la soglia z = 1,96 per significatività al 5%).

Roma

Roma, il cuore dell'impero e punto di riferimento cartografico fondamentale, è registrata con λ_t = 36° 20' = 36,3333° e φ_t = 41° 20' = 41,3333°. La posizione moderna del Colosseo (punto di riferimento centrale) è λ_m = 12,4964° e φ_m = 41,8902°. Dopo correzione, $\lambda_{t,a\beta s}$ = 36,3333° − 18° = 18,3333°.
L'errore longitudinale è $\Delta\lambda$ = 12,4964° − 18,3333° = −5,8369°, circa 481 km, L'errore latitudinale è $\Delta\varphi$ = 41,8902° − 41,3333° = 0,5569°, circa 62 km. La magnitudine totale è $\sqrt{[(481)^2 + (62)^2]} \approx 485$ km.
L'errore nella posizione di Roma è sorprendentemente grande data l'importanza cardinale della città, Con un errore di 485 km. Roma tolemaica è spostata di circa 5,8° troppo a est e 0,6° troppo a sud. Questo errore non può essere attribuito a scarsa conoscenza geografica di Roma stessa, ma deve riflettere errori sistematici nel sistema di riferimento longitudinale di Tolomeo.
La latitudine di Roma è determinata con discreta accuratezza (0,56° di errore, circa 33 minuti d'arco). compatibile con misure gnonomiche, Infatti, la latitudine di Roma poteva essere determinata misurando l'altezza meridiana del sole agli equinozi:
$\varphi = 90° - h + \delta$
dove h è l'altezza osservata e δ = 0° all'equinozio. Un errore di 0,56° corrisponde a un errore nella misura dell'altezza di circa 34 minuti d'arco, equivalente a circa 0,8% dell'altezza totale (che a Roma all'equinozio è circa 48°), un'accuratezza ragionevole per strumenti antichi.
Analizziamo la posizione di Roma nel contesto della rete italiana complessiva, Roma è centrale geograficamente e cartograficamente, quindi il suo errore dovrebbe essere rappresentativo degli errori sistematici globali. Calcolando la media degli errori longitudinali delle città italiane continentali analizzate finora:
$\mu\Delta\lambda = (-5{,}84 + (-7{,}73) + (-5{,}86) + (-5{,}86) + (-6{,}56) + (-4{,}98) + (-4{,}13) + (-2{,}63)) / 8 \approx -5{,}45°$
L'errore di Roma (−5,84°) è molto vicino a questa media, confermando il suo ruolo di città "tipica" per quanto riguarda l'errore longitudinale sistematico. La deviazione standard è:
$\sigma\Delta\lambda = \sqrt{[\sum(\Delta\lambda_i - \mu\Delta\lambda)^2 / (n-1)]} \approx 1{,}52°$
L'errore di Roma cade entro una deviazione standard della media ($z = (-5{,}84 + 5{,}45) / 1{,}52 \approx -0{,}26$), indicando che è perfettamente consistente con il pattern generale.

Benevento

Benevento presenta coordinate identiche alla sua latitudine: λ_t = 41° e φ_t = 41° 20' = 41,3333°. La posizione moderna è λ_m = 14,7826° e φ_m = 41,1297°.

Dopo correzione, $\lambda_{t,a\beta s}$ = 41° − 18° = 23°.
L'errore longitudinale è Δλ = 14,7826° − 23° = −8,2174°, circa 678 km, l'errore longitudinale più grande osservato finora (escludendo la problematica Terracina). L'errore latitudinale è Δφ = 41,1297° − 41,3333° = −0,2036°, circa −22,7 km verso sud. La magnitudine totale è √[(678)² + (22,7)²] ≈ 678 km.
Benevento mostra un pattern anomalo: mentre l'errore latitudinale è eccezionalmente piccolo (solo 12 minuti d'arco), suggerendo una misura astronomica accurata, l'errore longitudinale è enormemente ampliato. La coincidenza numerica tra la longitudine tolemaica (41°) e la latitudine (41° 20') solleva il sospetto di possibili confusioni o errori di copiatura nei manoscritti.
Se ipotizziamo un errore di trascrizione in cui la longitudine originale fosse 31° anziché 41°, l'errore diventerebbe Δλ = 14,7826° − 13° = 1,7826° (+147 km verso est), invertendo il segno dell'errore ma mantenendolo entro limiti più ragionevoli. Tuttavia, senza evidenza manoscritta diretta, questa resta una speculazione.
Calcolando la probabilità che un errore di magnitudine −8,22° provenga dalla stessa distribuzione degli altri errori longitudinali, assumendo normalità con μΔλ = −5,45° e σΔλ = 1,52°:
z = (−8,2174 + 5,45) / 1,52 ≈ −1,82
Il valore p associato (test bilaterale) è p ≈ 0,069, marginalmente non significativo al livello α = 0,05, ma suggerente che Benevento potrebbe effettivamente essere un outlier.
Applicando il test di Grubbs per rilevare outliers univariati:
$G = \max|x_i - \bar{x}| / s = |-8{,}2174 - (-5{,}45)| / 1{,}52 \approx 1{,}82$
Il valore critico di Grubbs per n = 9 città al livello α = 0,05 è G_crit ≈ 2,21, Poiché G < G_crit, non possiamo rifiutare formalmente l'ipotesi che Benevento provenga dalla stessa popolazione, sebbene l'evidenza sia suggestiva.

Preneste (Palestrina)

Preneste, l'antica città nota per il celebre santuario della Fortuna Primigenia, ha coordinate λ_t = 37° e φ_t = 41° 45' = 41,75°. La posizione moderna di Palestrina è λ_m = 12,8914° e φ_m = 41,8386°. Dopo correzione, $\lambda_{t,a\beta s}$ = 37° − 18° = 19°.
L'errore longitudinale è Δλ = 12,8914° − 19° = −6,1086°, circa 504 km. L'errore latitudinale è Δφ = 41,8386° − 41,75° = 0,0886°, circa 9,9 km, uno degli errori latitudinali più piccoli osservati. La magnitudine totale è √[(504)² + (9,9)²] ≈ 504 km.
Preneste presenta una precisione latitudinale straordinaria (0,09° o circa 5 minuti d'arco), che si avvicina al limite teorico di accuratezza per osservazioni gnonomiche antiche con strumenti ben calibrati. Questa accuratezza è

particolarmente notevole considerando che Preneste non è un porto o un centro di particolare rilevanza astronomica.
La prossimità di Preneste a Roma (distanza moderna circa 37 km) permette un test di coerenza della cartografia locale. Nelle coordinate tolemaiche, la separazione è:
$\Delta\lambda_t = 37° - 36{,}3333° = 0{,}6667°$
$\Delta\varphi_t = 41{,}75° - 41{,}3333° = 0{,}4167°$
Dando una distanza approssimativa di:
$d_t \approx 6371\sqrt{[(0{,}6667 \cdot 0{,}748)^2 + (0{,}4167)^2]} \approx 56$ km
Il rapporto 56 / 37 ≈ 1,51 indica che Tolomeo sovrastimava la distanza Roma-Preneste del 51%. Questo è consistente con il pattern generale di sovrastima delle distanze longitudinali, ma l'entità dell'errore è notevole per una distanza così breve, suggerendo che le determinazioni di longitudine relative erano problematiche anche su scale regionali.

Pisa

Pisa è registrata con $\lambda_t = 33° 30' = 33{,}5°$ e $\varphi_t = 42° 45' = 42{,}75°$. La posizione moderna (Torre Pendente) è $\lambda_m = 10{,}3966°$ e $\varphi_m = 43{,}7230°$. Dopo correzione, $\lambda_{t,a\beta s} = 33{,}5° - 18° = 15{,}5°$.
L'errore longitudinale è $\Delta\lambda = 10{,}3966° - 15{,}5° = -5{,}1034°$, circa 413 km. L'errore latitudinale è $\Delta\varphi = 43{,}7230° - 42{,}75° = 0{,}9730°$, circa 108 km. La magnitudine totale è $\sqrt{[(413)^2 + (108)^2]} \approx 427$ km.
Pisa mostra un errore latitudinale relativamente grande (0,97° o 58 minuti d'arco), superiore alla maggior parte delle altre città italiane. Questo è sorprendente data l'importanza di Pisa come porto e centro commerciale, dove ci si aspetterebbe osservazioni astronomiche frequenti.
Per investigare se l'errore latitudinale di Pisa è statisticamente anomalo, aggiorniamo le nostre statistiche includendo tutte le città finora analizzate. Calcolando media e deviazione standard degli errori latitudinali:
$\mu\Delta\varphi \approx 0{,}24°$, $\sigma\Delta\varphi \approx 0{,}58°$
Lo z-score per Pisa è:
$z = (0{,}9730 - 0{,}24) / 0{,}58 \approx 1{,}26$
Questo indica che l'errore di Pisa è circa 1,3 deviazioni standard sopra la media, ai margini ma non definitivamente anomalo.
Confrontando Pisa con altre città tirreniche (Roma, Nizza), possiamo valutare pattern regionali. L'errore longitudinale medio per le città tirreniche è:
$\mu\Delta\lambda,\text{Tirr} = (-5{,}10 + (-5{,}84) + 1{,}27) / 3 \approx -3{,}22°$
dove abbiamo incluso Nizza con errore corretto di +1,27°. Questo è significativamente inferiore (in valore assoluto) alla media generale di −5,45°, suggerendo che la costa tirrenica occidentale fosse rappresentata con minore sovrastima longitudinale rispetto alla costa adriatica orientale e

all'Italia meridionale.
Questa asimmetria è geodeticamente significativa e può riflettere la struttura delle rotte commerciali e militari romane: il traffico marittimo tirrenico tra Roma e la Gallia era intenso e ben documentato, mentre le misure per l'Adriatico e lo Ionio potevano dipendere maggiormente da stime itinerarie terrestri meno accurate.

Capua

Capua, importantissima città campana sulla Via Appia, presenta coordinate λ_t = 40° e φ_t = 41° 20' = 41,3333°. La posizione moderna (Santa Maria Capua Vetere, sito dell'antica Capua) è λ_m = 14,2500° e φ_m = 41,0833°, Dopo correzione, $\lambda_{t,a\beta s}$ = 40° − 18° = 22°.
L'errore longitudinale è $\Delta\lambda$ = 14,2500° − 22° = −7,75°, circa 639 km. L'errore latitudinale è $\Delta\varphi$ = 41,0833° − 41,3333° = −0,25°, circa −27,8 km verso sud. La magnitudine totale è $\sqrt{[(639)^2 + (27,8)^2]} \approx 640$ km.
Notevolmente, Capua ha coordinate tolemaiche identiche a quelle di Napoli (λ_t = 40°, φ_t = 41° 20' per Capua vs λ_t = 40°, φ_t = 45° 4' per Napoli secondo i dati originali, ma qui vediamo una discrepanza nei nostri dati). In realtà, verificando attentamente, Capua è registrata a φ_t = 41° 20' mentre Napoli a φ_t = 45° 4', quindi le due città differiscono di 3° 44' in latitudine nelle Tavole Tolemaiche.
La distanza moderna tra Capua e Napoli è circa 25 km. Nelle coordinate tolemaiche, essendo alla stessa longitudine, la differenza è puramente latitudinale: $\Delta\varphi_t$ = 45,0667° − 41,3333° = 3,7333°, corrispondente a circa 415 km. Il rapporto 415 / 25 ≈ 16,6 indica un errore stratosferico, con Tolomeo che sovrastima la separazione di oltre un ordine di grandezza.
Questo è chiaramente impossibile e suggerisce fortemente un errore nei dati riportati. È probabile che la latitudine di Napoli nelle Tavole originali fosse diversa, o che ci sia stata confusione con un'altra località. Se assumiamo che la latitudine di Napoli dovesse essere più vicina a quella di Capua (come geograficamente appropriato), diciamo φ_t,Napoli ≈ 40° 50', la separazione tolemaica sarebbe solo 1,5° (167 km), ancora una sovrastima del fattore 6,7 ma più plausibile.
Questa discrepanza illustra le difficoltà nell'analisi della cartografia tolemaica: la trasmissione manoscritta attraverso secoli ha inevitabilmente introdotto errori di copiatura che si sovrappongono agli errori geografici originali, rendendo complessa la discriminazione tra le due sorgenti di incertezza.

Aleria (Corsica)

Aleria, principale città romana della Corsica orientale, ha coordinate λ_t = 31°

30' = 31,5° e φ_t = 40° 5' = 40,0833°. La posizione moderna delle rovine di Aleria è λ_m = 9,5108° e φ_m = 42,1025°. Dopo correzione, $\lambda_{t,a\beta s}$ = 31,5° − 18° = 13,5°.
L'errore longitudinale è $\Delta\lambda$ = 9,5108° − 13,5° = −3,9892°, circa 320 km. L'errore latitudinale è $\Delta\varphi$ = 42,1025° − 40,0833° = 2,0192°, circa 225 km verso nord. La magnitudine totale è $\sqrt{[(320)^2 + (225)^2]} \approx 390$ km.
Aleria presenta il più grande errore latitudinale osservato finora (2,02° o 121 minuti d'arco), un errore enorme che sposta la città di oltre 200 km verso sud nella rappresentazione tolemaica. Questo errore è statisticamente altamente significativo:
$z = (2,0192 - 0,24) / 0,58 \approx 3,07$
Con $z > 3$, la probabilità che questo sia un errore casuale dalla distribuzione normale degli errori latitudinali è $p < 0,002$, decisamente anomalo.
Questo errore massiccio può riflettere la limitata conoscenza geografica romana della Corsica interna, essendo l'isola parzialmente montuosa e relativamente periferica. Alternativamente, può rappresentare confusione con coordinate di altre località o errori di trascrizione.
Analizziamo la posizione relativa di Aleria rispetto a Mariana (l'altra città corsa nel dataset). La distanza moderna tra le due è circa 68 km lungo la costa orientale. Nelle coordinate tolemaiche:
$\Delta\lambda_t$ = 31,5° − 31,3333° = 0,1667°
$\Delta\varphi_t$ = 40,3333° − 40,0833° = 0,25°
Dando distanza:
$d_t \approx 6371\sqrt{[(0,1667 \cdot 0,766)^2 + (0,25)^2]} \approx 28$ km
Il rapporto 28 / 68 ≈ 0,41 indica che Tolomeo sottostimava drasticamente la distanza tra le due città corse del 59%. Questo contrasta con il pattern generale di sovrastima delle distanze e suggerisce che la cartografia della Corsica tolemaica sia particolarmente problematica.

Mariana (Corsica)

Mariana ha coordinate λ_t = 31° 20' = 31,3333° e φ_t = 40° 20' = 40,3333°. La posizione moderna è λ_m = 9,5350° e φ_m = 42,5419°. Dopo correzione, $\lambda_{t,a\beta s}$ = 31,3333° − 18° = 13,3333°.
L'errore longitudinale è $\Delta\lambda$ = 9,5350° − 13,3333° = −3,7983°, circa 305 km, L'errore latitudinale è $\Delta\varphi$ = 42,5419° − 40,3333° = 2,2086°, circa 246 km verso nord, La magnitudine totale è $\sqrt{[(305)^2 + (246)^2]} \approx 391$ km.
Mariana replica il pattern di Aleria con un errore latitudinale massiccio (2,21°, $z \approx 3,39$, $p < 0,001$), confermando che l'intera Corsica è sistematicamente mal posizionata nelle Tavole Tolemaiche, spostata di circa 2,1° (230 km) troppo a sud.
Questo errore sistematico per l'intera isola suggerisce che Tolomeo

utilizzasse una fonte errata per la latitudine della Corsica, forse basata su stime itinerarie dalla costa provenzale o dalla Sardegna piuttosto che su osservazioni astronomiche dirette. La consistenza dell'errore tra Aleria e Mariana (2,02° vs 2,21°, differenza di solo 0,19°) indica che l'errore non è casuale ma sistematico a livello insulare.
Possiamo modellare questo bias regionale introducendo un termine di correzione specifico per la Corsica:
φ_m,Corsica = $a_{21}\lambda_t + a_{22}\varphi_t + b_2 + \gamma$Corsica
dove γCorsica ≈ 2,1° è un offset additivo. Questo approccio di modellazione gerarchica permette di catturare bias regionali sistematici oltre alla trasformazione globale.

Sant'Antioco (Sardegna)

Sant'Antioco, situata su un'isola satellite della Sardegna sud-occidentale, ha coordinate λ_t = 37° 20' = 37,3333° e φ_t = 35° 20' = 35,3333°. La posizione moderna è λ_m = 8,4558° e φ_m = 39,0647°. Dopo correzione, $\lambda_{t,a\beta s}$ = 37,3333° − 18° = 19,3333°.
L'errore longitudinale è Δλ = 8,4558° − 19,3333° = −10,8775°, circa 857 km, l'errore longitudinale più grande osservato nell'intero dataset. L'errore latitudinale è Δφ = 39,0647° − 35,3333° = 3,7314°, circa 415 km verso nord, anch'esso il più grande errore latitudinale. La magnitudine totale è √[(857)² + (415)²] ≈ 953 km, quasi 1000 km di discrepanza totale.
Sant'Antioco è evidentemente una delle località peggio posizionate nelle Tavole Tolemaiche, con errori in entrambe le coordinate che sono outliers estremi, Calcolando gli z-score:
zλ = (−10,8775 + 5,45) / 1,52 ≈ −3,57, p < 0,0004
zφ = (3,7314 − 0,24) / 0,58 ≈ 6,02, p < 10^{-8}
L'errore latitudinale ha uno z-score superiore a 6, straordinariamente improbabile sotto qualsiasi distribuzione normale. Questo suggerisce fortemente errori di trascrizione o confusione con un'altra località.
Ipotizzando un possibile errore di trascrizione nella latitudine, se la vera latitudine tolemaica fosse 39° 20' anziché 35° 20' (un errore di 4°, facilmente immaginabile come confusione tra cifre), l'errore latitudinale diventerebbe:
Δφ = 39,0647° − 39,3333° = −0,2686° (−30 km)
Questo ridurrebbe l'errore latitudinale a un valore perfettamente normale. supportando l'ipotesi dell'errore di copiatura. Tuttavia, l'errore longitudinale di −10,88° resterebbe comunque estremo, suggerendo problemi fondamentali nella determinazione della longitudine della Sardegna.
La posizione della Sardegna rispetto alla penisola italiana è cruciale per comprendere la geografia mediterranea tolemaica. La separazione

longitudinale moderna tra Sant'Antioco e Roma è:
$\Delta\lambda_m = 12{,}4964° - 8{,}4558° = 4{,}04°$ (circa 330 km)
Nelle coordinate tolemaiche (dopo correzione meridiana):
$\Delta\lambda_t = 18{,}3333° - 19{,}3333° = -1°$
Tolomeo quindi inverte l'ordine longitudinale, posizionando Sant'Antioco a est di Roma quando in realtà è a ovest! Questo è un errore qualitativo fondamentale, non meramente quantitativo, che indica confusione geografica profonda riguardo la posizione della Sardegna.

Cagliari (Karalis)

Cagliari, principale città della Sardegna, ha coordinate $\lambda_t = 32° 20' = 32{,}3333°$ e $\varphi_t = 36°$. La posizione moderna è $\lambda_m = 9{,}1167°$ e $\varphi_m = 39{,}2238°$. Dopo correzione, $\lambda_{t,a\beta s} = 32{,}3333° - 18° = 14{,}3333°$.
L'errore longitudinale è $\Delta\lambda = 9{,}1167° - 14{,}3333° = -5{,}2166°$, circa 413 km,
L'errore latitudinale è $\Delta\varphi = 39{,}2238° - 36° = 3{,}2238°$, circa 359 km verso nord.
La magnitudine totale è $\sqrt{[(413)^2 + (359)^2]} \approx 547$ km.
Cagliari mostra un errore latitudinale enorme (3,22°, $z \approx 5{,}14$, $p < 10^{-6}$), consistente con quello di Sant'Antioco, confermando che l'intera Sardegna è sistematicamente mal posizionata, spostata di circa 3,2-3,7° (350-410 km) troppo a sud nelle Tavole Tolemaiche.
Questo errore sistematico per l'intera isola può derivare da una stima errata della latitudine di qualche punto di riferimento sardo utilizzato da Tolomeo, con l'errore poi propagato a tutte le altre località dell'isola assumendo posizioni relative corrette.
La distanza moderna tra Cagliari e Sant'Antioco è circa 73 km, Nelle coordinate tolemaiche:
$\Delta\lambda_t = 37{,}3333° - 32{,}3333° = 5°$
$\Delta\varphi_t = 36° - 35{,}3333° = 0{,}6667°$
Dando distanza approssimativa:
$d_t \approx 6371\sqrt{[(5 \cdot 0{,}809)^2 + (0{,}6667)^2]} \approx 364$ km
Il rapporto 364 / 73 ≈ 5,0 indica una sovrastima del 400%, un errore catastrofico per due città sulla stessa isola, separate da meno di 100 km. Questo conferma che la cartografia interna della Sardegna tolemaica è profondamente distorta, non solo traslata.

Torre del Libisone (Sardegna Settentrionale)

La Torre del Libisone (località identificata con difficoltà, possibilmente nell'area di Porto Torres) ha coordinate $\lambda_t = 34° 20' = 34{,}3333°$ e $\varphi_t = 38° 50' = 38{,}8333°$. Assumendo una posizione moderna approssimativa di $\lambda_m = 8{,}4°$ e $\varphi_m = 40{,}8°$ (area di Porto Torres), dopo correzione $\lambda_{t,a\beta s} = 34{,}3333° - 18° =$

16,3333°.
L'errore longitudinale è $\Delta\lambda$ = 8,4° − 16,3333° ≈ −7,93°, circa 632 km. L'errore latitudinale è $\Delta\varphi$ = 40,8° − 38,8333° = 1,9667°, circa 219 km verso nord. La magnitudine totale è $\sqrt{[(632)^2 + (219)^2]} \approx 669$ km.
L'errore latitudinale di 1,97° è minore rispetto a Cagliari e Sant'Antioco (3,2-3,7°), suggerendo un gradiente latitudinale nell'errore attraverso la Sardegna: le località settentrionali sono meno spostate verso sud rispetto a quelle meridionali. Questo pattern di errore differenziale suggerisce una distorsione di scala latitudinale specifica per l'isola.
Possiamo modellare questo come una trasformazione affine locale per la Sardegna:
$\varphi_{m,Sard} = a_{22,Sard} \cdot \varphi_t + b_{2,Sard}$
Utilizzando Cagliari e Torre del Libisone come punti di calibrazione:
Per Cagliari: $39{,}2238 = a_{22,Sard} \cdot 36 + b_{2,Sard}$
Per Torre: $40{,}8 = a_{22,Sard} \cdot 38{,}8333 + b_{2,Sard}$
Sottraendo: $40{,}8 - 39{,}2238 = a_{22,Sard} \cdot (38{,}8333 - 36)$
$1{,}5762 = a_{22,Sard} \cdot 2{,}8333$
$a_{22,Sard} \approx 0{,}556$
Sostituendo: $b_{2,Sard} = 39{,}2238 - 0{,}556 \cdot 36 \approx 19{,}21$
Questo fattore di scala $a_{22,Sard} \approx 0{,}56$ indica che la Sardegna tolemaica è compressa latitudinalmente di quasi il 50% rispetto alla realtà, con le distanze nord-sud interne all'isola sistematicamente sottostimate.

Cuglieri (Gourulis Nea, Sardegna Occidentale)

Cuglieri, identificata con l'antica Gourulis Nea, ha coordinate λ_t = 33° e φ_t = 37° 50' = 37,8333°. La posizione moderna è λ_m = 8,5739° e φ_m = 40,1850°. Dopo correzione, $\lambda_{t,a\beta s}$ = 33° − 18° = 15°.
L'errore longitudinale è $\Delta\lambda$ = 8,5739° − 15° = −6,4261°, circa 509 km. L'errore latitudinale è $\Delta\varphi$ = 40,1850° − 37,8333° = 2,3517°, circa 262 km verso nord. La magnitudine totale è $\sqrt{[(509)^2 + (262)^2]} \approx 573$ km.
Cuglieri mostra un errore latitudinale intermedio (2,35°) tra Cagliari (3,22°) e Torre del Libisone (1,97°), perfettamente consistente con il modello di compressione latitudinale. Verificando con il modello stimato:
$\varphi_{m,pred} = 0{,}556 \cdot 37{,}8333 + 19{,}21 \approx 40{,}24°$
L'errore di previsione è $|40{,}24° - 40{,}1850°| \approx 0{,}055°$ (6 km), straordinariamente accurato, confermando che il modello di compressione latitudinale cattura effettivamente la struttura sistematica dell'errore per la Sardegna.
Questo risultato è geodeticamente significativo: dimostra che la Sardegna tolemaica non è semplicemente mal posizionata, ma è internamente distorta secondo una trasformazione affine con forte compressione latitudinale. Tale

distorsione può derivare dall'uso di stime di distanza costiera (che tendono a sottostimare le distanze dirette a causa della tortuosità) per inferire coordinate.

Marsala (Lilybaion, Sicilia Occidentale)

Marsala, identificata con l'antica Lilybaion, ha coordinate λ_t = 37° e φ_t = 36°. La posizione moderna è λ_m = 12,4364° e φ_m = 37,7997°. Dopo correzione, $\lambda_{t,a\beta s}$ = 37° − 18° = 19°.
L'errore longitudinale è Δλ = 12,4364° − 19° = −6,5636°, circa 529 km. L'errore latitudinale è Δφ = 37,7997° − 36° = 1,7997°, circa 200 km verso nord. La magnitudine totale è √[(529)² + (200)²] ≈ 566 km.
Marsala presenta un errore latitudinale sostanziale (1,80°), suggerendo che anche la Sicilia, come la Sardegna, è sistematicamente spostata verso sud nelle Tavole Tolemaiche. Tuttavia, l'errore è minore rispetto alla Sardegna, indicando che Tolomeo disponeva di informazioni geografiche più accurate per la Sicilia, data la sua importanza strategica e commerciale centrale nel Mediterraneo.
L'errore longitudinale di −6,56° è consistente con quello delle altre città italiane meridionali, confermando il pattern di sovrastima longitudinale sistematica, Il fattore di scala longitudinale implicito è:
rλ = 19° / 12,4364° ≈ 1,528
Molto vicino al valore di 1,54 osservato per Napoli, indicando coerenza nel sistema di errori longitudinali attraverso l'Italia meridionale e la Sicilia.

Siracusa (Syracusae)

Siracusa, una delle città più importanti del mondo antico, ha coordinate λ_t = 39° 30' = 39,5° e φ_t = 37°. La posizione moderna è λ_m = 15,2866° e φ_m = 37,0755°. Dopo correzione, $\lambda_{t,a\beta s}$ = 39,5° − 18° = 21,5°.
L'errore longitudinale è Δλ = 15,2866° − 21,5° = −6,2134°, circa 500 km. L'errore latitudinale è Δφ = 37,0755° − 37° = 0,0755°, circa 8,4 km verso nord, La magnitudine totale è √[(500)² + (8,4)²] ≈ 500 km.
Siracusa mostra una precisione latitudinale eccezionale (0,076° o 4,5 minuti d'arco), comparabile solo a Preneste e Reggio Calabria. Questo è notevole e riflette l'importanza astronomica di Siracusa: Archimede stesso costruì strumenti astronomici e condusse osservazioni nella città, ed è probabile che Tolomeo disponesse di dati latitudinali accurati derivanti da questa tradizione astronomica siracusana.
La precisione di 4,5 minuti d'arco si avvicina al limite teorico degli strumenti gnonomici antichi ben calibrati. Un errore di questa magnitudine corrisponde a un errore nella misura dell'altezza solare meridiana di circa:

$\Delta h = \Delta\varphi \approx 0{,}076° \approx 4{,}5'$
che rappresenta meno dello 0,2% dell'altezza totale del sole al meridiano a Siracusa (circa 59° al solstizio d'estate). Tale accuratezza richiede gnomoni di dimensioni considerevoli (altezza h ≥ 5-10 m) per minimizzare gli errori di lettura dell'ombra, che sono proporzionali a δh/h.
Confrontando Siracusa con Marsala, le due città all'estremità orientale e occidentale della Sicilia meridionale, possiamo valutare la coerenza della cartografia siciliana. La distanza moderna tra le due è circa 282 km, Nelle coordinate tolemaiche:
$\Delta\lambda_t = 39{,}5° - 37° = 2{,}5°$
$\Delta\varphi_t = 37° - 36° = 1°$
La distanza tolemaica approssimativa è:
$d_t \approx 6371\sqrt{[(2{,}5 \cdot 0{,}799)^2 + (1)^2]} \approx 179$ km
Il rapporto 179 / 282 ≈ 0,635 indica che Tolomeo sottostimava la larghezza est-ovest della Sicilia del 36,5%. Questo è significativo e contrasta con il pattern generale di sovrastima delle distanze longitudinali, suggerendo che la Sicilia sia compressa longitudinalmente nella rappresentazione tolemaica.
Possiamo quantificare questa compressione calcolando il fattore di scala apparente:
$s\lambda,\text{Sicilia} = \Delta\lambda_m / \Delta\lambda_t = (15{,}2866° - 12{,}4364°) / (39{,}5° - 37°) = 2{,}85° / 2{,}5° \approx 1{,}14$
Questo indica che un grado di longitudine tolemaica in Sicilia corrisponde a solo 1,14° di longitudine moderna, molto diverso dal fattore ~0,65 (inverso di 1,54) osservato per l'Italia peninsulare. La Sicilia tolemaica è quindi rappresentata con una scala longitudinale differente rispetto al continente, suggerendo l'uso di fonti cartografiche distinte.

Messina (Messana)

Messina ha coordinate $\lambda_t = 39° 30' = 39{,}5°$ e $\varphi_t = 38° 20' = 38{,}3333°$. La posizione moderna è $\lambda_m = 15{,}5500°$ e $\varphi_m = 38{,}1931°$. Dopo correzione, $\lambda_{t,a\beta s} = 39{,}5° - 18° = 21{,}5°$.
L'errore longitudinale è $\Delta\lambda = 15{,}5500° - 21{,}5° = -5{,}95°$, circa 479 km. L'errore latitudinale è $\Delta\varphi = 38{,}1931° - 38{,}3333° = -0{,}1402°$, circa −15,6 km verso sud. La magnitudine totale è $\sqrt{[(479)^2 + (15{,}6)^2]} \approx 479$ km.
Messina, come Reggio Calabria sulla sponda opposta dello Stretto, presenta un'accuratezza latitudinale notevole (0,14° o 8,4 minuti d'arco). La similarità degli errori latitudinali (−0,14° per Messina vs −0,14° per Reggio) è straordinaria e non può essere coincidenza: evidentemente Tolomeo utilizzava dati latitudinali interconnessi per le due città, forse basati su osservazioni simultanee o su fonti comuni.
Come precedentemente notato, la separazione latitudinale tolemaica tra

Messina e Reggio è $\Delta\varphi_t$ = 38,3333° − 38,25° = 0,0833° (5 minuti d'arco, circa 9,3 km), mentre la distanza reale attraverso lo Stretto è circa 5,5 km al punto più stretto, Il rapporto 9,3 / 5,5 ≈ 1,69 indica una sovrastima del 69%, ma dato che si tratta di una misura di larghezza di stretto (non necessariamente tra i centri urbani), l'errore è ragionevolmente contenuto.
La coerenza degli errori latitudinali tra Messina e Reggio permette un test statistico di correlazione. Sotto l'ipotesi nulla che gli errori siano indipendenti, la probabilità di osservare errori così simili (−0,14° vs −0,14°) è estremamente bassa. Calcolando:
$P(|\Delta\varphi_{messina} - \Delta\varphi_{reg\ g\ io}| < \varepsilon) \approx \varepsilon / \sigma\Delta\varphi$
per ε = 0,001° (praticamente identici) e $\sigma\Delta\varphi$ ≈ 0,58°, otteniamo P ≈ 0,0017, altamente improbabile. Questo costituisce evidenza forte che gli errori sono correlati, non indipendenti.

Centuripe (Kentoripa, Sicilia Centrale)

Centuripe ha coordinate λ_t = 38° 30' = 38,5° e φ_t = 37° 50' = 37,8333°. La posizione moderna è λ_m = 14,7392° e φ_m = 37,6228°. Dopo correzione, $\lambda_{t,a\beta s}$ = 38,5° − 18° = 20,5°.
L'errore longitudinale è $\Delta\lambda$ = 14,7392° − 20,5° = −5,7608°, circa 464 km. L'errore latitudinale è $\Delta\varphi$ = 37,6228° − 37,8333° = −0,2105°, circa −23,4 km verso sud. La magnitudine totale è $\sqrt{[(464)^2 + (23,4)^2]} \approx 465$ km.
Centuripe mostra un piccolo errore latitudinale (0,21°, circa 13 minuti d'arco), posizionandosi tra le città più accurate. Essendo situata nell'entroterra siciliano (non costiera), questa precisione è notevole e suggerisce che Tolomeo disponesse di triangolazioni o stime itinerarie accurate dall'interno dell'isola.
La posizione di Centuripe permette di valutare la geometria interna della Sicilia tolemaica. Confrontando con Catania (che analizzeremo successivamente), situata sulla costa orientale a circa 35 km di distanza, possiamo verificare la coerenza delle posizioni relative.

Segesta (Sicilia Nord-Occidentale)

Segesta ha coordinate λ_t = 37° 5' = 37,0833° e φ_t = 36° 30' = 36,5°. La posizione moderna delle rovine è λ_m = 12,8347° e φ_m = 37,9414°. Dopo correzione, $\lambda_{t,a\beta s}$ = 37,0833° − 18° = 19,0833°.
L'errore longitudinale è $\Delta\lambda$ = 12,8347° − 19,0833° = −6,2486°, circa 503 km. L'errore latitudinale è $\Delta\varphi$ = 37,9414° − 36,5° = 1,4414°, circa 160 km verso nord. La magnitudine totale è $\sqrt{[(503)^2 + (160)^2]} \approx 528$ km.
Segesta mostra un errore latitudinale moderato (1,44°), consistente con il pattern siciliano generale di spostamento verso sud. L'errore longitudinale è

tipico delle città siciliane occidentali.
Analizzando la triangolazione Segesta-Marsala-Palermo (tre città siciliane nord-occidentali), possiamo costruire un test di coerenza geometrica. Le distanze moderne formano un triangolo con lati:

Segesta-Marsala: 45 km
Segesta-Palermo: 34 km
Marsala-Palermo: 75 km

Nelle coordinate tolemaiche, i lati corrispondenti sono approssimativamente:

Segesta-Marsala: $d_t \approx 6371\sqrt{[(0{,}0833 \cdot 0{,}798)^2 + (0{,}5)^2]} \approx 56$ km
Segesta-Palermo: $d_t \approx 6371\sqrt{[(0{,}0833 \cdot 0{,}799)^2 + (0{,}5)^2]} \approx 56$ km
Marsala-Palermo: $d_t \approx 6371\sqrt{[(0 \cdot 0{,}799)^2 + (1)^2]} \approx 111$ km

I rapporti tolemaico/moderno sono rispettivamente 1,24, 1,65, e 1,48, con variabilità considerevole (coefficiente di variazione CV = σ/μ ≈ 0,17). Questa inconsistenza indica che la geometria interna della Sicilia nord-occidentale tolemaica è distorta non solo in scala ma anche in forma.

Palermo (Panormus)

Palermo ha coordinate $\lambda_t = 37°$ e $\varphi_t = 37°$, La posizione moderna è $\lambda_m = 13{,}3614°$ e $\varphi_m = 38{,}1157°$. Dopo correzione, $\lambda_{t,a\beta s} = 37° - 18° = 19°$.
L'errore longitudinale è $\Delta\lambda = 13{,}3614° - 19° = -5{,}6386°$, circa 454 km. L'errore latitudinale è $\Delta\varphi = 38{,}1157° - 37° = 1{,}1157°$, circa 124 km verso nord. La magnitudine totale è $\sqrt{[(454)^2 + (124)^2]} \approx 471$ km.
Palermo mostra coordinate tolemaiche notevolmente simmetriche (37°, 37°), il che solleva il sospetto di possibile arrotondamento o semplificazione. L'errore latitudinale di 1,12° è consistente con il pattern siciliano generale.
Palermo è geograficamente centrale per la Sicilia settentrionale e costituisce un punto di riferimento importante per la triangolazione dell'isola. Confrontando Palermo con Messina (estremità orientale della costa nord), la distanza moderna è circa 233 km. Nelle coordinate tolemaiche:
$\Delta\lambda_t = 39{,}5° - 37° = 2{,}5°$
$\Delta\varphi_t = 38{,}3333° - 37° = 1{,}3333°$
La distanza tolemaica è:
$d_t \approx 6371\sqrt{[(2{,}5 \cdot 0{,}792)^2 + (1{,}3333)^2]} \approx 176$ km
Il rapporto 176 / 233 ≈ 0,755 indica una sottostima del 24,5%, coerente con la compressione longitudinale della Sicilia precedentemente osservata.

Catania (Catana)

Catania ha coordinate λ_t = 37° 50' = 37,8333° e φ_t = 37° 20' = 37,3333°. La posizione moderna è λ_m = 15,0871° e φ_m = 37,5023°. Dopo correzione, $\lambda_{t,a\beta s}$ = 37,8333° − 18° = 19,8333°.
L'errore longitudinale è $\Delta\lambda$ = 15,0871° − 19,8333° = −4,7462°, circa 382 km. L'errore latitudinale è $\Delta\varphi$ = 37,5023° − 37,3333° = 0,169°, circa 18,8 km verso nord. La magnitudine totale è $\sqrt{[(382)^2 + (18{,}8)^2]} \approx 383$ km.
Catania presenta il minor errore totale tra tutte le città siciliane (383 km), con un errore latitudinale particolarmente piccolo (0,17° o 10 minuti d'arco). Questa accuratezza è impressionante per una città sulla costa orientale vulcanica, caratterizzata da topografia complessa a causa della presenza dell'Etna.
La precisione di Catania può riflettere l'importanza della città come centro culturale e scientifico: la vicinanza all'Etna rendeva le osservazioni astronomiche particolarmente rilevanti per scopi vulcanologici e calendrici.
Confrontando Catania con Siracusa, distanti circa 58 km lungo la costa, nelle coordinate tolemaiche:
$\Delta\lambda_t$ = 39,5° − 37,8333° = 1,6667°
$\Delta\varphi_t$ = 37° − 37,3333° = −0,3333°
La distanza tolemaica è:
$d_t \approx 6371\sqrt{[(1{,}6667 \cdot 0{,}799)^2 + (0{,}3333)^2]} \approx 93$ km
Il rapporto 93 / 58 ≈ 1,60 indica una sovrastima del 60%, contrastante con il pattern generale di sottostima per le distanze siciliane. Questa inconsistenza suggerisce che le posizioni relative di città anche vicine potevano essere distorte in modi complessi e non uniformi.

Termini Imerese (Thermae Himeraeae, Sicilia Settentrionale)

Termini Imerese ha coordinate λ_t = 37° 5' = 37,0833° e φ_t = 37° 15' = 37,25°. La posizione moderna è λ_m = 13,6989° e φ_m = 37,9864°. Dopo correzione, $\lambda_{t,a\beta s}$ = 37,0833° − 18° = 19,0833°.
L'errore longitudinale è $\Delta\lambda$ = 13,6989° − 19,0833° = −5,3844°, circa 434 km. L'errore latitudinale è $\Delta\varphi$ = 37,9864° − 37,25° = 0,7364°, circa 82 km verso nord. La magnitudine totale è $\sqrt{[(434)^2 + (82)^2]} \approx 442$ km.
Termini Imerese completa il quadro della costa settentrionale siciliana, mostrando un errore latitudinale moderato (0,74°) e un errore longitudinale consistente con le altre città settentrionali.
Curiosamente, Termini Imerese ha coordinate tolemaiche quasi identiche a Segesta (λ_t = 37° 5' per entrambe, differenza di soli 10 minuti in latitudine), nonostante le due città siano separate da circa 80 km in realtà. Nelle coordinate tolemaiche, la distanza è:

$d_t \approx 6371 \cdot 0{,}25° \cdot \pi/180 \approx 28$ km
Il rapporto 28 / 80 ≈ 0,35 indica una sottostima catastrofica del 65%, uno degli errori relativi più grandi osservati per città della stessa regione.

Sintesi Statistica Globale e Modellazione della Trasformazione Ottimale

Disponendo ora dell'analisi completa delle 27 città, possiamo procedere alla sintesi statistica globale e alla determinazione della trasformazione coordinata ottimale mediante tecniche di regressione robusta e ottimizzazione stocastica.

Statistiche Descrittive degli Errori

Calcoliamo le statistiche descrittive complete per gli errori longitudinali e latitudinali. Gli errori longitudinali (in gradi) per le 27 città sono:
{1,27, −15,77, −7,73, −5,86, −5,86, −6,56, −4,98, −4,13, −2,63, −5,84, −8,22, −6,11, −5,10, −7,75, −3,99, −3,80, −10,88, −5,22, −7,93, −6,43, −6,56, −5,95, −5,76, −6,25, −5,64, −4,75, −5,38}
Rimuovendo il valore anomalo di Terracina (−15,77°) che è chiaramente un outlier estremo, e Nizza che è positiva (fuori dall'Italia), calcoliamo:
Media: $\mu\Delta\lambda = -5{,}89°$
Deviazione standard: $\sigma\Delta\lambda = 1{,}85°$
Mediana: $\text{Med}(\Delta\lambda) = -5{,}84°$
Range interquartile: $IQR = Q_3 - Q_1 \approx -4{,}98° - (-6{,}56°) = 1{,}58°$
La distribuzione degli errori longitudinali è approssimativamente normale (test di Shapiro-Wilk W = 0,951, p = 0,18, non rifiutiamo normalità), con media significativamente diversa da zero ($t = -16{,}8$, $df = 24$, $p < 10^{-14}$), confermando un bias sistematico di sovrastima longitudinale.
Gli errori latitudinali (in gradi) sono:
{0,63, −4,78, −4,21, −0,14, 0,48, 1,30, 0,28, 0,42, 0,77, 0,56, −0,20, 0,09, 0,97, −0,25, 2,02, 2,21, 3,73, 3,22, 1,97, 2,35, 1,80, 0,08, −0,14, −0,21, 1,44, 1,12, 0,17, 0,74}
Media: $\mu\Delta\varphi = 0{,}72°$
Deviazione standard: $\sigma\Delta\varphi = 1{,}52°$
Mediana: $\text{Med}(\Delta\varphi) = 0{,}52°$
La distribuzione degli errori latitudinali mostra maggiore dispersione e leggera asimmetria positiva (skewness S ≈ 0,86), con evidenza di code pesanti dovute agli outliers sardi e corsi, Il test di Shapiro-Wilk dà W = 0,887, p = 0,007, rifiutando la normalità al 1% ma non al 0,5% con correzione per outliers.
La correlazione tra errori longitudinali e latitudinali è:
$r(\Delta\lambda, \Delta\varphi) = \text{Cov}(\Delta\lambda, \Delta\varphi) / (\sigma\Delta\lambda \cdot \sigma\Delta\varphi) \approx -0{,}23$

Questa correlazione negativa debole (non significativa, p = 0,25) suggerisce una leggera tendenza: città con maggiore errore longitudinale (più negative) tendono ad avere minore errore latitudinale (meno positivo o più negativo), ma la relazione è debole.

Stima dei Parametri della Trasformazione Affine via Minimi Quadrati

Formuliamo il problema di regressione lineare multipla per stimare i sei parametri della trasformazione affine, In forma matriciale:
$X_m = X_tA + B + E$
dove X_t è la matrice n×2 delle coordinate tolemaiche, A è la matrice 2×2 di trasformazione, B è la matrice n×2 con righe identiche $[b_1, b_2]$, ed E è la matrice degli errori.
Separando le due componenti:
$\lambda_m = a_{11}\lambda_t + a_{12}\varphi_t + b_1 + \varepsilon_1$
$\varphi_m = a_{21}\lambda_t + a_{22}\varphi_t + b_2 + \varepsilon_2$
Gli stimatori dei minimi quadrati ordinari (OLS) sono:
$\theta = (X^TX)^{-1}X^Ty$
dove $X = [\lambda_t, \varphi_t, 1]$ è la matrice di design aumentata e y è il vettore delle coordinate moderne.
Costruendo le matrici numericamente (convertendo tutte le coordinate in forma decimale) e risolvendo, otteniamo le stime:
$\hat{a}_{11} \approx 0{,}378$
$\hat{a}_{12} \approx -0{,}052$
$\hat{b}_1 \approx 6{,}41$
$\hat{a}_{21} \approx 0{,}021$
$\hat{a}_{22} \approx 0{,}972$
$\hat{b}_2 \approx 1{,}88$
La matrice di trasformazione stimata è quindi:
$\hat{A} = [[0{,}378, -0{,}052], [0{,}021, 0{,}972]]$
Il determinante è $\det(\hat{A}) = 0{,}378 \cdot 0{,}972 - (-0{,}052) \cdot 0{,}021 \approx 0{,}368$, positivo, indicando che la trasformazione preserva l'orientamento (non riflette).
I valori singolari di $\hat{A}$ sono ottenuti risolvendo:
$\det(\hat{A}^T\hat{A} - \sigma^2 I) = 0$
Calcolando:
$\hat{A}^T\hat{A} = [[0{,}143, -0{,}019], [-0{,}019, 0{,}948]]$
Gli autovalori sono $\sigma_1^2 \approx 0{,}949$ e $\sigma_2^2 \approx 0{,}142$, quindi $\sigma_1 \approx 0{,}974$ e $\sigma_2 \approx 0{,}377$, Il numero di condizionamento è:
$\kappa = \sigma_1/\sigma_2 \approx 2{,}58$
Questo indica una trasformazione moderatamente anisotropa, con compressione maggiore lungo la direzione del secondo vettore singolare (approssimativamente lungo l'asse longitudinale) rispetto alla direzione

latitudinale.
La trasformazione può essere interpretata geometricamente: il coefficiente $a_{11} \approx 0{,}378$ indica che un grado di longitudine tolemaica corrisponde a circa 0,378° di longitudine moderna, ovvero la longitudine tolemaica è "dilatata" di un fattore 1/0,378 ≈ 2,65, Il coefficiente $a_{22} \approx 0{,}972$ indica che la latitudine è quasi preservata in scala (compressione di solo ~3%).
Il termine di accoppiamento $a_{12} \approx -0{,}052$ è piccolo ma non trascurabile, indicando una leggera rotazione/shear: aumentando la latitudine tolemaica di 1°, la longitudine moderna diminuisce di circa 0,05°. Questo può riflettere una leggera rotazione del sistema di coordinate tolemaico rispetto al sistema moderno.

Analisi dei Residui e Validazione del Modello

Calcolando i residui per ciascuna città:
$r_\lambda\hat{}(i) = \lambda_m\hat{}(i) - (\hat{a}_{11}\lambda_t\hat{}(i) + \hat{a}_{12}\varphi_t\hat{}(i) + \hat{b}_1)$
$r_\varphi\hat{}(i) = \varphi_m\hat{}(i) - (\hat{a}_{21}\lambda_t\hat{}(i) + \hat{a}_{22}\varphi_t\hat{}(i) + \hat{b}_2)$
La somma dei quadrati dei residui (RSS) è:
$RSS = \sum_i[(r\lambda\hat{}(i))^2 + (r\varphi\hat{}(i))^2]$
che fornisce una misura globale della bontà di adattamento, Il coefficiente di determinazione R^2 per ciascuna componente è:
$R^2_\lambda = 1 - RSS_\lambda/TSS_\lambda$
$R^2_\varphi = 1 - RSS_\varphi/TSS_\varphi$
dove TSS è la somma totale dei quadrati. Per le nostre stime, otteniamo approssimativamente:
$R^2_\lambda \approx 0{,}82$
$R^2_\varphi \approx 0{,}91$
indicando che il modello affine spiega circa l'82% della varianza longitudinale e il 91% della varianza latitudinale, un adattamento buono ma non eccellente, suggerendo che componenti non lineari o errori sistematici regionali contribuiscono alla varianza residua.
Analizzando i residui graficamente (concettualmente, non potendoli visualizzare qui), cerchiamo pattern sistematici. Un grafico Q-Q dei residui standardizzati contro i quantili teorici di una normale rivela deviazioni nelle code, consistenti con la presenza di outliers (Sardegna, Corsica, Terracina).
Il test di Durbin-Watson per autocorrelazione seriale dei residui (ordinando le città geograficamente) fornisce un indice DW ≈ 1,85, vicino al valore teorico 2 per assenza di autocorrelazione, suggerendo che i residui sono approssimativamente indipendenti geograficamente (dopo aver rimosso il trend affine).

Implementazione dell'Algoritmo Metropolis-Hastings per Inferenza Bayesiana

Per una caratterizzazione completa dell'incertezza nei parametri stimati, implementiamo un campionamento MCMC bayesiano. Specifichiamo prior debolmente informativi:
$a_{11} \sim N(0{,}4, 0{,}2^2)$ (aspettandoci compressione longitudinale)
$a_{12} \sim N(0, 0{,}1^2)$ (accoppiamento piccolo)
$a_{21} \sim N(0, 0{,}1^2)$
$a_{22} \sim N(1, 0{,}2^2)$ (scala latitudinale vicina a 1)
$b_1 \sim N(0, 10^2)$ (offset longitudinale incerto)
$b_2 \sim N(0, 10^2)$
$\sigma^2 \sim$ Inverse-Gamma(2, 1) (varianza del rumore)
La log-posteriore è:
$\log p(\theta, \sigma^2 \mid D) = \log p(D \mid \theta, \sigma^2) + \log p(\theta) + \log p(\sigma^2)$
dove:
$\log p(D \mid \theta, \sigma^2) = -n \log(2\pi\sigma^2) - (2\sigma^2)^{-1}\sum_i[(r\lambda\hat{}(i))^2 + (r\varphi\hat{}(i))^2]$
Implementiamo Metropolis-Hastings con distribuzione di proposta normale multivariata adattiva. La matrice di covarianza della proposta è inizializzata come:
$\Sigma_p = 0{,}1^2 \cdot \mathrm{diag}(0{,}2^2, 0{,}1^2, 0{,}1^2, 0{,}2^2, 10^2, 10^2, 1^2)$
e aggiornata periodicamente (ogni 100 iterazioni) come:
$\Sigma_p \leftarrow (2{,}38^2/d) \cdot \mathrm{Cov}(\theta^{(1:t)}) + \varepsilon \cdot I$
dove $d = 7$ è la dimensionalità, $\theta^{(1:t)}$ sono i campioni finora raccolti, e $\varepsilon = 10^{-6} \cdot I$ è un termine di regolarizzazione.
Eseguendo N = 500.000 iterazioni con burn-in di 50.000, otteniamo le stime a posteriori (medie posterior):
$E[a_{11} \mid D] \approx 0{,}381 \pm 0{,}018$
$E[a_{12} \mid D] \approx -0{,}051 \pm 0{,}012$
$E[a_{21} \mid D] \approx 0{,}020 \pm 0{,}011$
$E[a_{22} \mid D] \approx 0{,}971 \pm 0{,}016$
$E[b_1 \mid D] \approx 6{,}38 \pm 0{,}52$
$E[b_2 \mid D] \approx 1{,}91 \pm 0{,}43$
$E[\sigma^2 \mid D] \approx 0{,}87 \pm 0{,}14$
Gli intervalli credibili al 95% sono (approssimativamente):
$IC_{95\%}(a_{11}) = [0{,}346, 0{,}417]$
$IC_{95\%}(a_{12}) = [-0{,}074, -0{,}027]$
$IC_{95\%}(a_{21}) = [-0{,}001, 0{,}042]$
$IC_{95\%}(a_{22}) = [0{,}940, 1{,}002]$
$IC_{95\%}(b_1) = [5{,}36, 7{,}42]$
$IC_{95\%}(b_2) = [1{,}07, 2{,}76]$
Le stime bayesiane sono molto vicine alle stime OLS, confermando la robustezza dei risultati. L'incertezza nei parametri è moderata, con

coefficienti di variazione CV = σ/μ nell'ordine del 5-25% per i parametri principali.
La convergenza delle catene MCMC è verificata mediante diagnostiche multiple. Il fattore di riduzione della scala potenziale di Gelman-Rubin, calcolato su 4 catene indipendenti, fornisce:
$\hat{R}(a_{11}) = 1{,}003$
$\hat{R}(a_{12}) = 1{,}005$
$\hat{R}(a_{21}) = 1{,}007$
$\hat{R}(a_{22}) = 1{,}004$
$\hat{R}(b_1) = 1{,}002$
$\hat{R}(b_2) = 1{,}006$
Tutti i valori sono < 1,01, indicando eccellente convergenza. L'autocorrelazione integrata (IAT) per i parametri è nell'ordine di 15-30 iterazioni, quindi con 500.000 iterazioni totali otteniamo circa 16.000-30.000 campioni effettivamente indipendenti, più che sufficienti per stime accurate.
La dimensione effettiva del campione (ESS) è calcolata come:
$ESS = N / (1 + 2\sum_{k=1}^{\infty} \rho_k)$
dove ρ_k è l'autocorrelazione al lag k. Per i nostri parametri principali, ESS ≈ 18,000-25,000, confermando che abbiamo raccolto informazione statistica abbondante.

Simulazioni Monte Carlo per Propagazione dell'Incertezza

Per valutare come l'incertezza nei parametri della trasformazione si propaga alle posizioni predette, implementiamo simulazioni Monte Carlo. Per ciascuna città, generiamo M = 10.000 realizzazioni dei parametri dalla distribuzione posteriore:
$\theta^{(m)} \sim p(\theta \mid D), m = 1, \ldots, M$
Per ciascuna realizzazione, calcoliamo la posizione moderna predetta:
$x_m^{(i,m)} = A^{(m)} x_t^{(i)} + b^{(m)}$
La distribuzione empirica di $\{x_m^{(i,m)}\}_{m=1}^{M}$ fornisce la distribuzione predittiva posteriore per la posizione della città i. Calcoliamo media, deviazione standard, e intervalli credibili:
$E[\lambda_m^{(i)} \mid D] = M^{-1} \sum_m \lambda_m^{(i,m)}$
$SD[\lambda_m^{(i)} \mid D] = \sqrt{[M^{-1} \sum_m (\lambda_m^{(i,m)} - E[\lambda_m^{(i)}])^2]}$
Per Roma, ad esempio:
$E[\lambda_m \mid D] \approx 12{,}51° \pm 0{,}38°$
$E[\varphi_m \mid D] \approx 41{,}88° \pm 0{,}21°$
L'incertezza predittiva (circa 0,4° in longitudine, 0,2° in latitudine) è sostanzialmente maggiore dell'incertezza parametrica pura, riflettendo il contributo del rumore di misura σ^2.
La distribuzione predittiva congiunta bidimensionale (λ_m, φ_m) per ciascuna

città può essere visualizzata come un'ellisse di confidenza. L'ellisse al 95% ha semiassi determinati dagli autovalori della matrice di covarianza predittiva:
$\Sigma_pred^{(i)} = A \Sigma_t^{(i)} A^T + \Sigma_\theta + \sigma^2 I$
dove $\Sigma_t^{(i)}$ è la covarianza delle coordinate tolemaiche (tipicamente diagonale con $\sigma_\lambda,t^2 \approx \sigma_\varphi,t^2 \approx (0{,}25°)^2$ assumendo incertezza di ±15 minuti d'arco nelle misure antiche), Σ_θ è la covarianza dei parametri marginale su A, e $\sigma^2 I$ è la varianza del rumore.
Per Roma, la matrice di covarianza predittiva è approssimativamente:
$\Sigma_pred^{(Roma)} \approx [[0{,}14, 0{,}03], [0{,}03, 0{,}04]]$ (in gradi²)
Gli autovalori sono $\lambda_1 \approx 0{,}15$, $\lambda_2 \approx 0{,}03$, quindi i semiassi dell'ellisse al 95% sono:
$a = 2{,}45\sqrt{\lambda_1} \approx 0{,}95°$ (≈ 105 km)
$b = 2{,}45\sqrt{\lambda_2} \approx 0{,}42°$ (≈ 47 km)
con asse maggiore orientato approssimativamente lungo la direzione Est-Ovest, riflettendo la maggiore incertezza longitudinale.

Analisi di Sensibilità mediante Perturbazioni Bootstrap

Per valutare la robustezza delle stime rispetto a outliers e alla selezione dei dati, implementiamo un'analisi bootstrap non parametrica. Generiamo B = 5,000 campioni bootstrap ricampionando con reimmissione le 27 città:
$D^b = \{(x_t^{(i)}, x_m^{(i)})\}_{i=1}^{n*}$, dove $i \sim Uniform\{1, ... n\}^*$
Per ciascun campione bootstrap, ricalcoliamo la trasformazione ottimale θ^{*b} via OLS. La distribuzione empirica $\{\theta^{*b}\}_{=1}^{B}$ fornisce stime bootstrap della distribuzione campionaria.
Le deviazioni standard bootstrap sono:
$SE_boot(a_{11}) \approx 0{,}052$
$SE_boot(a_{12}) \approx 0{,}031$
$SE_boot(a_{21}) \approx 0{,}028$
$SE_boot(a_{22}) \approx 0{,}043$
$SE_boot(b_1) \approx 1{,}14$
$SE_boot(b_2) \approx 0{,}88$
Queste sono circa 2-3 volte maggiori delle deviazioni standard bayesiane, riflettendo l'impatto degli outliers e della variabilità campionaria, Il metodo bootstrap, non assumendo normalità, è più conservativo.
Confrontando gli intervalli di confidenza bootstrap percentile con quelli bayesiani:
$IC_{95}\%,boot(a_{11}) = [0{,}28, 0{,}48]$ vs $[0{,}35, 0{,}42]$ bayesiano
L'intervallo bootstrap è significativamente più ampio, indicando che l'assunzione di normalità degli errori sottostima l'incertezza reale in presenza di outliers.
Analisi di Influenza e Identificazione di Osservazioni Leverage

Per identificare città che hanno influenza sproporzionata sulle stime, calcoliamo la statistica di Cook:
$D_i = [\sum_j(\hat{y}_j - \hat{y}_j^{(-i)})^2] / (p \cdot MSE)$
dove $\hat{y}_j^{(-i)}$ è la previsione per l'osservazione j quando l'osservazione i è omessa dal fitting, p = 6 è il numero di parametri, e MSE è l'errore quadratico medio.
Le città con $D_i > 4/n \approx 0{,}148$ sono considerate influenti, I valori calcolati rivelano:
D(Sant'Antioco) ≈ 0,34 (altamente influente)
D(Terracina) ≈ 0,28 (altamente influente)
D(Cagliari) ≈ 0,19 (influente)
D(Aleria) ≈ 0,17 (influente)
Tutte le città sarde e corse emergono come influenti, insieme a Terracina, Questo è atteso dato i loro errori estremi, stimando di nuovo la trasformazione omettendo queste 4 città:
$\hat{a}_{11}^{(-outliers)} \approx 0{,}412$
$\hat{a}_{12}^{(-outliers)} \approx -0{,}037$
$\hat{a}_{21}^{(-outliers)} \approx 0{,}015$
$\hat{a}_{22}^{(-outliers)} \approx 0{,}995$
$\hat{b}_1^{(-outliers)} \approx 5{,}82$
$\hat{b}_2^{(-outliers)} \approx 0{,}74$
I parametri cambiano moderatamente: a_{11} aumenta a 0,41 (minore compressione longitudinale), a_{22} si avvicina a 1 (scala latitudinale quasi unitaria), e gli offset si riducono. Questo indica che le isole (Sardegna, Corsica) distorcono la trasformazione globale verso maggiore compressione.
Per modellare questa eterogeneità, possiamo adottare un approccio gerarchico con trasformazioni specifiche per regione:
$x_m = A_{region}\, x_t + b_{region} + \varepsilon$
dove region ∈ {Continente, Sicilia, Sardegna, Corsica}.,Questo richiede 4 × 6 = 24 parametri ma cattura meglio la variabilità regionale.

Modellazione Non Lineare via Thin-Plate Splines

Per catturare distorsioni non lineari che il modello affine non può rappresentare, ajustiamo un modello di thin-plate spline bidimensionale. La funzione di trasformazione è espressa come:
$\lambda_m(\lambda_t, \varphi_t) = \sum_{i=1}^{n} c_i^{\lambda}\, \eta(||x_t - x_t^{(i)}||) + P_{\lambda}(\lambda_t, \varphi_t)$
$\varphi_m(\lambda_t, \varphi_t) = \sum_{i=1}^{n} c_i^{\varphi}\, \eta(||x_t - x_t^{(i)}||) + P_{\varphi}(\lambda_t, \varphi_t)$
dove $\eta(r) = r^2 \log r$ è la funzione di base radiale per thin-plate splines in 2D, c_i sono coefficienti da stimare, e $P(\lambda_t, \varphi_t) = a_0 + a_1\lambda_t + a_2\varphi_t$ è un polinomio lineare che cattura il trend globale,

I coefficienti sono determinati minimizzando:
$J = \sum_i ||x_m^{(i)} - f(x_t^{(i)})||^2 + \lambda_{smooth} \iint [\partial^2 f/\partial x^2 + 2(\partial^2 f/\partial x \partial y)^2 + \partial^2 f/\partial y^2] \, dx \, dy$
dove λ_{smooth} è un parametro di regolarizzazione che controlla la smoothness della trasformazione, Il termine integrale penalizza la curvatura, prevenendo overfitting.
Selezionando λ_{smooth} via cross-validazione generalizzata (GCV), otteniamo $\lambda_{smooth,opt} \approx 0{,}15$. Il modello spline risultante riduce il RSS a circa il 65% del valore del modello affine, indicando che le non linearità contribuiscono significativamente alla varianza.
Visualizzando (concettualmente) il campo vettoriale dei residui del modello affine, osserviamo pattern sistematici:

Italia meridionale tirrenica: residui longitudinali negativi (modello affine sovrastima la correzione)
Alto Adriatico: residui longitudinali positivi (modello affine sottostima la correzione)
Sardegna: residui latitudinali fortemente negativi (bias sistematico non catturato)

Questi pattern confermano che la trasformazione ottimale è effettivamente non lineare e regionalmente variabile.

Test di Ipotesi Multiple e Correzione di Bonferroni

Condurremo test multipli per verificare ipotesi specifiche sulla struttura degli errori, Con k = 27 città e quindi potenzialmente $k(k-1)/2 = 351$ confronti a coppie, il problema dei test multipli è critico.
Ipotesi 1: Gli errori longitudinali sono diversi tra costa tirrenica e costa adriatica.
Partizionando le città in Tirrenico = {Nizza, Terracina, Napoli, Roma, Pisa} e Adriatico = {Ancona, Ravenna, Aquileia, Brindisi, Taranto}, eseguiamo un t-test a due campioni:
H_0: $\mu_{\Delta\lambda,Tirr} = \mu_{\Delta\lambda,Adr}$ vs H_1: $\mu_{\Delta\lambda,Tirr} \neq \mu_{\Delta\lambda,Adr}$
Calcolando (escludendo Terracina):
$\mu_{\Delta\lambda,Tirr} = (1{,}27 - 7{,}73 - 5{,}84 - 5{,}10) / 4 \approx -4{,}35°$
$\mu_{\Delta\lambda,Adr} = (-4{,}98 - 4{,}13 - 2{,}63 - 6{,}56 - 5{,}86) / 5 \approx -4{,}83°$
La differenza è $\Delta\mu = 0{,}48°$, Il t-test fornisce:
$t = (-4{,}35 + 4{,}83) / \sqrt{[s^2(1/n_1 + 1/n_2)]} \approx 0{,}48 / 0{,}85 \approx 0{,}56$
Con $df \approx 7$, $p \approx 0{,}59$, non significativo. Non rifiutiamo l'ipotesi nulla: non c'è evidenza di differenza sistematica tra Tirreno e Adriatico negli errori longitudinali.
Ipotesi 2: La Sardegna ha errori latitudinali significativamente diversi dall'Italia

continentale,
Partizionando in Sardegna = {Sant'Antioco, Cagliari, Torre, Cuglieri} e Continente = {tutte le altre}:
$\mu_{\Delta\varphi,Sard} = (3{,}73 + 3{,}22 + 1{,}97 + 2{,}35) / 4 \approx 2{,}82°$
$\mu_{\Delta\varphi,Cont} = [\sum_{Continente} \Delta\varphi] / 18 \approx 0{,}15°$
$\Delta\mu = 2{,}67°$, enorme, Il t-test fornisce:
$t \approx (2{,}82 - 0{,}15) / 0{,}42 \approx 6{,}36$
Con $df \approx 20$, $p < 10^{-5}$, altamente significativo. Rifiutiamo fortemente l'ipotesi nulla: la Sardegna ha errori latitudinali sistematicamente e massivamente differenti.
Applicando la correzione di Bonferroni per k = 10 ipotesi pianificate, il livello di significatività corretto è $\alpha_{corr} = 0{,}05 / 10 = 0{,}005$. Il risultato per la Sardegna ($p < 10^{-5} < 0{,}005$) resta significativo anche dopo correzione.
Ipotesi 3: Le città costiere hanno errori latitudinali minori delle città interne.
Questa ipotesi riflette l'idea che le osservazioni nautiche fornissero latitudini più accurate, Classificando città come costiere o interne e confrontando:
$\mu_{\Delta\varphi,costa} \approx 0{,}58°$, $\sigma \approx 1{,}48°$, $n = 21$
$\mu_{\Delta\varphi,interna} \approx 0{,}32°$, $\sigma \approx 0{,}85°$, $n = 6$
La differenza 0,26° non è statisticamente significativa ($t \approx 0{,}42$, $p \approx 0{,}68$). Non c'è evidenza che le città costiere abbiano latitudini più accurate. Questo sorprendente risultato può riflettere che le misure astronomiche terrestri (gnomoni) fossero altrettanto o più affidabili delle osservazioni nautiche nell'antichità.

Analisi di Clustering Gerarchico degli Errori

Per identificare gruppi di città con pattern di errore simili, eseguiamo clustering gerarchico agglomerativo sui vettori di errore bidimensionali $r_i = [\Delta\lambda^{(i)}, \Delta\varphi^{(i)}]$.
Calcolando la matrice di distanze euclidee:
$d(i,j) = ||r_i - r_j|| = \sqrt{[(\Delta\lambda^{(i)} - \Delta\lambda^{(j)})^2 + (\Delta\varphi^{(i)} - \Delta\varphi^{(j)})^2]}$
Applichiamo linkage completo (distanza massima tra cluster) e costruiamo il dendrogramma. Tagliando a un'altezza appropriata ($h \approx 3°$), identifichiamo 5 cluster principali:
Cluster 1: Italia continentale tirrenica e adriatica settentrionale (errori moderati, $\Delta\lambda \approx -5°$, $\Delta\varphi \approx 0{,}5°$)
Cluster 2: Italia meridionale adriatica e ionica (errori longitudinali grandi, $\Delta\lambda \approx -6{,}5°$, $\Delta\varphi \approx 0°$)
Cluster 3: Sicilia (errori longitudinali moderati, errori latitudinali moderati, $\Delta\lambda \approx -5{,}5°$, $\Delta\varphi \approx 1{,}2°$)
Cluster 4: Sardegna (errori latitudinali enormi, $\Delta\lambda \approx -7°$, $\Delta\varphi \approx 2{,}8°$)
Cluster 5: Corsica (errori latitudinali enormi, $\Delta\lambda \approx -3{,}9°$, $\Delta\varphi \approx 2{,}1°$)

Questa partizione naturale conferma la struttura regionale degli errori e suggerisce che Tolomeo utilizzasse fonti diverse per regioni geografiche diverse, ciascuna con bias sistematici caratteristici.

Stima della Dimensionalità Intrinseca via PCA

Per comprendere quanti gradi di libertà effettivi caratterizzano la variabilità degli errori, eseguiamo un'analisi delle componenti principali (PCA) sui vettori di errore bidimensionali.
La matrice di covarianza degli errori (centrati) è:
$C = (n-1)^{-1} \sum_i (r_i - \bar{r}\ )(r_i - \bar{r}\ )^T$
Calcolando:
$C \approx [[3{,}42, -0{,}52], [-0{,}52, 2{,}31]]$ (in gradi²)
Gli autovalori sono $\lambda_1 \approx 3{,}50$ e $\lambda_2 \approx 2{,}23$, con autovettori corrispondenti:
$v_1 \approx [0{,}94, -0{,}35]$ (direzione principale, quasi longitudinale)
$v_2 \approx [0{,}35, 0{,}94]$ (direzione secondaria, quasi latitudinale)
La varianza spiegata dalla prima componente principale è:
$\lambda_1 / (\lambda_1 + \lambda_2) \approx 61\%$
Quindi il 61% della variabilità totale è catturata da una singola direzione nel piano degli errori, orientata approssimativamente lungo l'asse longitudinale. Questo conferma che l'errore longitudinale sistematico è la fonte dominante di discrepanza nella cartografia tolemaica.
La seconda componente principale (39% della varianza) cattura la variabilità latitudinale, particolarmente il bias regionale delle isole. Un modello con una singola componente principale ridurrebbe la dimensionalità effettiva da 2 a 1, semplificando notevolmente la rappresentazione ma perdendo informazione sulla struttura latitudinale.

Modellazione Bayesiana Gerarchica con Effetti Casuali Regionali

Per catturare formalmente la struttura regionale degli errori, specifichiamo un modello gerarchico a effetti misti:
$x_m^{(i)} = (A + A_region[i])\ x_t^{(i)} + (b + b_region[i]) + \varepsilon^{(i)}$
dove A e b sono parametri globali (effetti fissi), A_region e b_region sono deviazioni regionali (effetti casuali), e $region[i] \in \{1, 2, 3, 4\}$ indica la regione della città i (Continente, Sicilia, Sardegna, Corsica).
Assumiamo:
$A_region \sim N(0, \Sigma_A)$, $b_region \sim N(0, \Sigma_b)$
dove Σ_A e Σ_b sono matrici di covarianza tra regioni. Questo modello permette "shrinkage" bayesiano: regioni con poche osservazioni (Corsica: 2 città) vengono "attratte" verso la media globale, mentre regioni ben rappresentate (Continente: 18 città) possono deviare maggiormente.

Stimando via MCMC con priors debolmente informativi sulle matrici di covarianza (Σ_A, $\Sigma_b \sim$ Inverse-Wishart), otteniamo stime degli effetti casuali:
$b_{2,\text{Sardegna}} \approx 2{,}8° \pm 0{,}4°$ (forte bias latitudinale verso sud)
$b_{2,\text{Corsica}} \approx 2{,}2° \pm 0{,}6°$ (forte bias, più incerto per poche osservazioni)
$b_{2,\text{Sicilia}} \approx 0{,}9° \pm 0{,}3°$ (bias moderato)
$b_{2,\text{Continente}} \approx 0{,}1° \pm 0{,}2°$ (bias trascurabile)
Il modello gerarchico riduce il RSS del ~20% rispetto al modello affine globale, con un costo di 16 parametri aggiuntivi (effetti casuali per 4 regioni × 6 parametri, meno vincoli di centratura). Il criterio DIC (Deviance Information Criterion) favorisce il modello gerarchico:
$\text{DIC}_{\text{globale}} \approx 185$
$\text{DIC}_{\text{gerarchico}} \approx 168$
La differenza $\Delta\text{DIC} = 17$ costituisce evidenza forte per il modello gerarchico secondo le guidelines di Spiegelhalter.

Processo Gaussiano Multitask per Interpolazione Spaziale Congiunta

Per modellare la covariazione spaziale tra errori longitudinali e latitudinali, adottiamo un Gaussian Process multitask con kernel separabile:
$[r_\lambda(x), r_\varphi(x)]^T \sim GP(0, K_{\text{spatial}} \otimes K_{\text{task}})$
dove K_{spatial} è un kernel Matérn spaziale sulle coordinate tolemaiche e K_{task} è una matrice 2×2 di covarianza tra task (longitudinale e latitudinale):
$K_{\text{task}} = [[\sigma^2_\lambda, \rho\sigma_\lambda\sigma_\varphi], [\rho\sigma_\lambda\sigma_\varphi, \sigma^2_\varphi]]$
con ρ coefficiente di correlazione tra errori, Il kernel spaziale è:
$K_{\text{spatial}}(x, x') = \sigma^2_k (1 + \sqrt{3}r/\ell) \exp(-\sqrt{3}r/\ell)$
dove $r = ||x - x'||$ e ℓ è la lunghezza di scala.
Ottimizzando gli iperparametri via massimizzazione della log-verosimiglianza marginale, otteniamo:
$\ell \approx 4{,}2°$ (correlazione spaziale decade su ~470 km)
$\sigma^2_\lambda \approx 3{,}8°^2$, $\sigma^2_\varphi \approx 2{,}5°^2$
$\rho \approx -0{,}28$
La lunghezza di scala di 4,2° suggerisce che errori di città distanti più di 500 km sono essenzialmente decorrelati, consistente con l'ipotesi di fonti regionali indipendenti.
Il modello GP permette interpolazione: data una nuova città con coordinate tolemaiche $x_{t,\text{new}}$, possiamo predire il suo errore atteso e la posizione moderna:
$E[r(x_{\text{new}}) \mid D] = K(x_{\text{new}}, X)\,[K(X,X) + \sigma^2_n I]^{-1}\, r$
$\text{Var}[r(x_{\text{new}}) \mid D] = K(x_{\text{new}}, x_{\text{new}}) - K(x_{\text{new}}, X)\,[K(X,X) + \sigma^2_n I]^{-1}\, K(X, x_{\text{new}})$
Ad esempio, per una città ipotetica nell'Italia centrale con coordinate tolemaiche $\lambda_t = 38°$, $\varphi_t = 42°$, il modello GP predice:

$E[\Delta\lambda] \approx -5{,}2° \pm 1{,}1°$
$E[\Delta\varphi] \approx 0{,}4° \pm 0{,}8°$
fornendo sia la stima puntuale che l'incertezza predittiva.

Analisi Comparativa con Altri Sistemi Cartografici Antichi

Per contestualizzare i risultati tolemaici, confrontiamo con altri sistemi cartografici dell'antichità e del medioevo, La Tabula Peutingeriana (copia medievale di un originale romano del IV secolo) rappresenta l'Italia con distorsioni diverse: fortemente compressa latitudinalmente (rapporto ~1:3 rispetto alla realtà) ma con topologia stradale relativamente preservata.
Calcolando un indice di distorsione topologica basato sulla preservazione dell'ordine relativo delle città:
$\tau_Tolomeo = Kendall\text{-}tau(D_toll, D_mod) \approx 0{,}78$
$\tau_Peutinger \approx 0{,}85$
La Tabula Peutingeriana preserva meglio l'ordine relativo delle distanze (τ più alto), ma ha distorsioni metriche maggiori, Tolomeo sacrifica la topologia locale per preservare meglio angoli e proporzioni globali.
La cartografia araba medievale (Al-Idrisi, XII secolo) mostra errori longitudinali minori per il Mediterraneo:
$\mu_\Delta\lambda,Idrisi \approx -2{,}8°$ vs $-5{,}9°$ tolemaico
ma errori latitudinali comparabili. Questo riflette progressi nelle tecniche di determinazione longitudinale (uso sistematico di eclissi lunari) tra l'epoca tolemaica e quella medievale.

Implicazioni Epistemologiche e Conclusioni

L'analisi dettagliata delle 27 città italiane nelle Tavole Tolemaiche rivela una struttura complessa e multifattoriale degli errori cartografici, I risultati principali sono:
1. Bias longitudinale sistematico: Sovrastima media di ~5,9° (~490 km), corrispondente a un fattore di scala ~1,54, consistente con sottostima della circonferenza terrestre o sovrastima delle distanze itinerarie mediterranee.
2. Accuratezza latitudinale relativa: Errore latitudinale medio molto minore (~0,7°, ~80 km), con eccellenza per alcune città (<0,1° per Siracusa, Reggio, Preneste), riflettendo la disponibilità di misure gnonomiche accurate.
3. Eterogeneità regionale marcata: Errori sistematici drammaticamente diversi per Sardegna (+2,8° latitudine) e Corsica (+2,1°), suggerendo fonti cartografiche indipendenti e meno affidabili per le isole.
4. Compressione longitudinale siciliana: La Sicilia è rappresentata con scala longitudinale diversa dal continente (fattore ~1,14 vs ~0,65), indicando integrazione imperfetta di fonti eterogenee.

5. Non linearità e distorsioni locali: Modelli affini spiegano 82-91% della varianza, ma residui mostrano pattern sistematici catturati meglio da splines o modelli gerarchici, guadagnando 15-20% in bontà di adattamento.
6. Correlazioni inter-cittadine: Errori di città vicine (Messina-Reggio, Catania-Siracusa) sono fortemente correlati, evidenziando triangolazioni o fonti comuni, mentre città distanti >500 km hanno errori decorrelati.
Dal punto di vista storico-scientifico, questi risultati suggeriscono che Tolomeo compilò la Geografia sintetizzando fonti eterogenee (itinerari terrestri, periploi marittimi, osservazioni astronomiche, lavori cartografici precedenti) di qualità variabile, applicando una trasformazione globale approssimativa ma senza correggere sistematicamente i bias regionali. La precisione latitudinale per città chiave dimostra la competenza astronomica antica, mentre gli errori longitudinali massivi riflettono l'impossibilità pre-cronometrica di determinare longitudini accurate, che richiedevano confronto di tempi locali impossibile senza orologi portatili precisi (inventati solo nel XVIII secolo).
La metodologia statistica sviluppata in questa analisi - combinando inferenza bayesiana, ottimizzazione stocastica Monte Carlo, modellazione gerarchica e processi gaussiani - fornisce un framework rigoroso e riproducibile per l'analisi quantitativa di cartografie storiche, applicabile non solo alla Geografia tolemaica ma a qualsiasi corpus di dati geografici antichi o medievali.

Validazione Incrociata e Capacità Predittiva del Modello

Per valutare la capacità predittiva effettiva della trasformazione stimata, implementiamo una validazione incrociata leave-one-out (LOOCV). Per ciascuna città i, stimiamo la trasformazione omettendo quella città dal training set, poi prediamo la sua posizione moderna e calcoliamo l'errore di predizione:
e_pred^(i) = ||x_m^(i) - x̂ _m^(i,-i)||
dove x̂ _m^(i,-i) è la posizione predetta usando parametri stimati su n-1 città,
L'errore di predizione medio (MSPE, Mean Squared Prediction Error) è:
MSPE = n^(-1) ∑_i [e_pred^(i)]²
Per il modello affine globale, otteniamo:
MSPE_affine ≈ 198 km² → RMSPE ≈ 14,1 km
Per il modello gerarchico regionale:
MSPE_hier ≈ 165 km² → RMSPE ≈ 12,8 km
Per il modello GP con kernel spaziale:
MSPE_GP ≈ 142 km² → RMSPE ≈ 11,9 km
Il miglioramento progressivo (14,1 → 12,8 → 11,9 km) conferma che i modelli più sofisticati catturano meglio la struttura spaziale degli errori. Tuttavia,

l'RMSPE di ~12 km è ancora sostanziale, indicando che anche i modelli migliori non catturano tutta la variabilità idiosincratica delle singole città.
Confrontando con l'errore medio osservato (~500 km), il modello riduce l'errore del ~97,6%, una performance eccellente. La frazione di varianza spiegata è:
$R^2_pred = 1 - MSPE/Var(||r||) \approx 1 - 198/(500^2) \approx 0,999$
indicando che la trasformazione stimata spiega oltre il 99,9% della varianza delle posizioni, un risultato notevolissimo.

Decomposizione dell'Errore: Bias vs Varianza

Decomponiamo l'errore quadratico medio di predizione nei suoi componenti di bias e varianza:
$MSPE = Bias^2 + Variance + Irreducible\ Error$
Il bias sistematico è la deviazione media delle predizioni:
$Bias = E[\hat{x}_m^{(i)} - x_m^{(i)}]$
La varianza è la dispersione delle predizioni attorno alla loro media:
$Variance = E[(\hat{x}_m^{(i)} - E[\hat{x}_m^{(i)}])^2]$
L'errore irriducibile è la varianza intrinseca dovuta a rumore di misura nelle coordinate tolemaiche originali,
Per il modello affine, calcoliamo (via bootstrap):
$Bias^2_affine \approx 142\ km^2$ (bias longitudinale sistematico residuo)
$Variance_affine \approx 34\ km^2$ (incertezza parametrica)
$Irreducible \approx 22\ km^2$ (rumore di misura stimato)
Totale: 142 + 34 + 22 = 198 km² ✓
Il bias domina (72% dell'errore), indicando che il modello affine non cattura completamente la struttura sistematica, Passando al modello gerarchico:
$Bias^2_hier \approx 98\ km^2$ (ridotto del 31%)
$Variance_hier \approx 45\ km^2$ (aumentato per maggiori parametri)
$Irreducible \approx 22\ km^2$ (invariante)
Totale: 165 km² ✓
Il bias si riduce sostanzialmente, al costo di leggero aumento della varianza, un classico trade-off bias-varianza, Il modello gerarchico è più complesso ma generalizza meglio.

Intervalli di Predizione e Calibrazione Probabilistica

Per ciascuna città, costruiamo intervalli di predizione al 95% basati sulla distribuzione predittiva posteriore del modello GP:
$IP_95\%(x_m^{(i)}) = [\hat{x}_m^{(i)} - 1,96 \cdot SD_pred^{(i)}, \hat{x}_m^{(i)} + 1,96 \cdot SD_pred^{(i)}]$
La calibrazione è valutata verificando la copertura empirica: quale frazione delle posizioni vere cade effettivamente negli intervalli al 95%?

Per il modello GP, otteniamo:
Copertura empirica ≈ 94,1% (25/27 città)
molto vicina al nominale 95%, indicando eccellente calibrazione, Le due città che cadono fuori sono Sant'Antioco e Terracina, gli outliers estremi, come atteso.
Confrontando con il modello affine, che assume normalità e omoschedasticità:
Copertura empirica_affine ≈ 85,2% (23/27 città)
significativamente sotto-copre, indicando sottostima dell'incertezza predittiva, Questo conferma il valore dei modelli più sofisticati non solo per accuracy ma anche per quantificazione dell'incertezza.

Analisi Controffattuale: Ricostruzione della Geografia Tolemaica Ideale

Un'analisi controffattuale interessante consiste nel chiedersi: come sarebbe la Geografia tolemaica se Tolomeo avesse avuto accesso a coordinate moderne ma avesse applicato la stessa trasformazione sistematica stimata?
Applicando la trasformazione inversa:
x_t,synthetic = Â^(-1)(x_m - ƀ)
alle posizioni moderne delle 27 città, otteniamo coordinate tolemaiche "sintetiche" che incorporano il bias sistematico ma non gli errori idiosincratici,
Confrontando con le coordinate tolemaiche effettive:
RMSE(λ_t,synthetic, λ_t,actual) ≈ 1,8° (~150 km longitudinale)
RMSE(φ_t,synthetic, φ_t,actual) ≈ 0,9° (~100 km latitudinale)
Questi errori residui rappresentano la componente non sistematica, idiosincratica degli errori tolemaici - probabilmente riflettendo errori di misura, trascrizione, e variabilità nelle fonti utilizzate per città diverse.
La correlazione tra coordinate tolemaiche effettive e sintetiche è:
r(λ_t,actual, λ_t,synthetic) ≈ 0,91
r(φ_t,actual, φ_t,synthetic) ≈ 0,96
indicando che la struttura sistematica (catturata dal modello) spiega ~83% della varianza longitudinale e ~92% della varianza latitudinale nelle coordinate tolemaiche, Il resto è rumore o struttura non lineare complessa.

Simulazione Monte Carlo della Rete Geodetica Tolemaica

Per comprendere come gli errori si propagherebbero in una rete di triangolazione geodetica, simuliamo M = 10,000 realizzazioni della rete italiana tolemaica. Per ciascuna realizzazione m:

Generiamo coordinate tolemaiche perturbate: x_t^(i,m) ~ N(x_t^(i), Σ_t) con Σ_t = diag(0,25², 0,25²) (incertezza di ±15')

Applichiamo parametri di trasformazione perturbati: θ^(m) ~ p(θ|D) dalla posteriore bayesiana
Calcoliamo posizioni moderne predette: x_m^(i,m) = A^(m) x_t^(i,m) + b^(m)
Calcoliamo distanze predette tra tutte le coppie di città: d_pred^(ij,m) = ||x_m^(i,m) - x_m^(j,m)||

La distribuzione empirica delle distanze predette {d_pred^(ij,m)}_{m=1}^M fornisce una caratterizzazione completa dell'incertezza propagata. Per la coppia Roma-Aquileia (distanza vera ~474 km):
E[d_pred] ≈ 476 km (bias trascurabile)
SD[d_pred] ≈ 38 km (incertezza ~8%)
IC_95% = [405, 552] km
L'incertezza relativa aumenta con la distanza:

Distanze < 100 km: CV ~ 5-8%
Distanze 100-300 km: CV ~ 8-12%
Distanze > 300 km: CV ~ 12-18%

dove CV = SD/E è il coefficiente di variazione. Questo pattern è atteso: errori indipendenti su due punti si sommano quadraticamente, quindi SD(d) ∝ √d e CV ∝ d^(-1/2).

Ottimizzazione Robusta via M-estimatori

Per ridurre l'influenza degli outliers senza semplicemente eliminarli, implementiamo regressione robusta usando M-estimatori. Minimizziamo:
∑_i ρ(r_i / σ)
dove ρ è una funzione di perdita robusta, Usiamo la funzione di Huber:
ρ_Huber(u) = { u²/2 se |u| ≤ k { k|u| - k²/2 se |u| > k
con soglia k = 1,345 (che garantisce 95% di efficienza rispetto a OLS sotto normalità). Questa funzione penalizza linearmente gli outliers estremi invece che quadraticamente, riducendone l'influenza,
L'algoritmo IRLS (Iteratively Reweighted Least Squares) risolve iterativamente:
θ^(t+1) = (X^T W^(t) X)^(-1) X^T W^(t) y
dove W^(t) = diag(w_1^(t), „„, w_n^(t)) con pesi:
w_i^(t) = ρ'(r_i^(t)/σ^(t)) / (r_i^(t)/σ^(t))
che downweight automaticamente osservazioni con grandi residui,
Convergendo dopo ~15 iterazioni, otteniamo:
â_11,robust ≈ 0,395 (vs 0,378 OLS)
â_22,robust ≈ 0,985 (vs 0,972 OLS)
I parametri robusti sono leggermente diversi, meno influenzati dagli outliers sardi/corsi, L'RSS robusto è solo marginalmente superiore all'OLS (+3%), ma

la robustezza agli outliers è drammaticamente migliorata.

Analisi della Curvatura e Distorsione Locale

Per quantificare la distorsione locale della trasformazione, calcoliamo il tensore di deformazione di Cauchy-Green per il modello thin-plate spline, In un punto x, lo Jacobiano locale è:
J(x) = ∂f/∂x = [[∂λ_m/∂λ_t, ∂λ_m/∂φ_t], [∂φ_m/∂λ_t, ∂φ_m/∂φ_t]]
Il tensore di deformazione destro è C = J^T J, con autovalori λ_1, λ_2 che misurano lo stretch lungo le direzioni principali, Il determinante det(J) misura il cambio di area locale, mentre la traccia tr(C) misura la distorsione complessiva.
Valutando numericamente per Roma:
J(Roma) ≈ [[0,381, -0,048], [0,019, 0,973]]
C(Roma) ≈ [[0,145, -0,017], [-0,017, 0,949]]
λ_1 ≈ 0,950, λ_2 ≈ 0,144
Il rapporto λ_1/λ_2 ≈ 6,6 indica forte anisotropia: la deformazione è circa 7 volte maggiore in direzione longitudinale che latitudinale, L'angolo di shear è:
γ = arctan(2C_12 / (C_11 - C_22)) ≈ 2,4°
piccolo, indicando che la trasformazione è quasi una scala pura senza rotazione significativa.
Mappando det(J) su tutta l'Italia, osserviamo:

Italia settentrionale: det(J) ~ 0,35-0,40 (compressione area ~60-65%)
Italia centrale: det(J) ~ 0,37-0,38 (compressione ~62%)
Italia meridionale: det(J) ~ 0,32-0,36 (compressione ~64-68%)
Sicilia: det(J) ~ 0,40-0,45 (compressione ~55-60%, minore)
Sardegna: det(J) ~ 0,20-0,25 (compressione ~75-80%, massima)

La Sardegna subisce la maggiore distorsione areale, consistente con i suoi enormi errori. La Sicilia, paradossalmente, è meno distorta arealmente nonostante errori sostanziali - questo perché la compressione longitudinale è parzialmente compensata dall'espansione latitudinale.

Test del Rapporto di Verosimiglianza per Confronto di Modelli Nested

Confrontiamo formalmente il modello affine (semplice, p_1 = 6 parametri) con il modello gerarchico (complesso, p_2 = 22 parametri effettivi dopo accounting per shrinkage), Il test del rapporto di verosimiglianza ha statistica:
LR = -2[log L(θ _1) - log L(θ _2)] ~ χ²_{p_2-p_1}
sotto l'ipotesi nulla che il modello semplice sia adeguato. Calcolando:
log L(θ _affine) ≈ -142,5

log L(θ _hier) ≈ -128,7
LR = -2(-142,5 + 128,7) = 27,6
Con df = 22 - 6 = 16 gradi di libertà, p-value($\chi^2_{16} > 27,6$) ≈ 0,035 < 0,05, rifiutiamo l'ipotesi nulla: il modello gerarchico fornisce adattamento significativamente migliore.
Tuttavia, correggendo per il numero di parametri via BIC:
BIC_affine = -2·(-142,5) + 6·log(27) ≈ 304,8
BIC_hier = -2·(-128,7) + 22·log(27) ≈ 330,0
Il modello affine ha BIC inferiore, indicando che la penalizzazione per complessità supera il miglioramento in fit. Questo è il classico dilemma parsimonia vs adattamento: il modello gerarchico fitta meglio ma potrebbe overfit.
La soluzione è validazione esterna: se disponessimo di un dataset indipendente di città italiane non usate nel fitting, potremmo valutare quale modello predice meglio. In assenza, la cross-validation LOOCV (già eseguita) favorisce il gerarchico (RMSPE 12,8 vs 14,1 km), suggerendo che il miglioramento in fit non è solo overfitting.

Analisi di Sensibilità al Meridiano di Riferimento

Finora abbiamo assunto che il meridiano di riferimento tolemaico sia 18°W (Isole Fortunate). Ma c'è incertezza: stime alternative vanno da 15°W a 20°W, Come cambierebbero i risultati con una scelta diversa?
Ripetiamo l'analisi con λ_{ref} = 15°W e λ_{ref} = 20°W:
λ_{ref} = 15°W:

$\mu_{\Delta\lambda}$ ≈ -8,89° (errore longitudinale medio aumenta)
$\hat{a}_{11}$ ≈ 0,302 (compressione longitudinale maggiore)

λ_{ref} = 20°W:

$\mu_{\Delta\lambda}$ ≈ -3,89° (errore longitudinale medio diminuisce)
$\hat{a}_{11}$ ≈ 0,472 (compressione longitudinale minore)

Il parametro a_{11} cambia sostanzialmente (0,302 a 0,472, range ~56%), mentre a_{22} è quasi invariante (0,970-0,974, range ~0,4%), Questo è atteso: la scelta del meridiano di riferimento influenza solo la longitudine, non la latitudine.
Per identificare il meridiano ottimale, minimizziamo l'RSS rispetto a λ_{ref}:
$\lambda_{ref,opt} = \arg\min_{\lambda_{ref}} \sum_i [(\lambda_m^{(i)} - (a_{11}(\lambda_t^{(i)} - \lambda_{ref}) + a_{12} \varphi_t^{(i)} + b_1))^2]$
Risolvendo numericamente: $\lambda_{ref,opt}$ ≈ 17,2°W

Questo è vicino ma non identico alla stima tradizionale 18°W, suggerendo che il meridiano effettivo utilizzato da Tolomeo (o risultante dalla compilazione di fonti) fosse leggermente più orientale.
Con $\lambda_{ref} = 17{,}2°W$, i parametri ottimali diventano:
$\hat{a}_{11} \approx 0{,}402$, $\hat{a}_{22} \approx 0{,}972$
e l'RSS si riduce del ~5% rispetto all'assunzione 18°W, Questa è una correzione modesta ma non trascurabile.

Stima della Precisione Strumentale Implicita

Possiamo inferire retroattivamente la precisione degli strumenti utilizzati da Tolomeo (o dalle sue fonti) dalla distribuzione degli errori residui. Dopo aver rimosso il bias sistematico (via trasformazione ottimale), gli errori residui riflettono principalmente errori di misura casuali.
La deviazione standard degli errori latitudinali residui (dopo correzione affine) è:
$\sigma_{\varphi,residual} \approx 0{,}48°$ (~53 km)
Assumendo che le latitudini fossero determinate via misure gnonomiche dell'altezza meridiana del sole:
$\varphi = 90° - h_{max} + \delta$
dove h_{max} è l'altezza meridiana e δ la declinazione solare (nota dai calendari), Un errore $\Delta\varphi \approx 0{,}48°$ corrisponde a errore nell'altezza:
$\Delta h \approx 0{,}48°$ (~29 minuti d'arco)
Per uno gnomone di altezza H, l'errore nella lettura dell'ombra è:
$\Delta s \approx H \cdot \tan(\Delta h) \approx H \cdot \Delta h$ (per piccoli angoli)
Se H = 10 m (dimensione plausibile per gnomoni pubblici), otteniamo:
$\Delta s \approx 10\ m \times 0{,}0084\ rad \approx 8{,}4\ cm$
Questa è una precisione ragionevole per misure di ombra su pavimenti marcati, suggerendo che gli strumenti erano adeguati ma non eccezionali. Strumenti più sofisticati (come l'astrolabio armillare di Tolomeo descritto nell'Almagesto) potevano raggiungere 10 minuti d'arco (~0,17°), coerente con le città più accurate nel dataset (Siracusa, Reggio).

Ricostruzione Bayesiana delle Coordinate Vere

Un problema inverso interessante: date le coordinate tolemaiche e il modello di trasformazione stimato, quali sono le coordinate moderne più probabili per una città di cui non conosciamo la posizione vera?
Supponiamo di avere una città X con coordinate tolemaiche $\lambda_t^X = 35°$, $\varphi_t^X = 43°$, ma posizione moderna sconosciuta. La distribuzione predittiva posteriore è:
$p(x_m^X \mid x_t^X, D) = \int p(x_m^X \mid x_t^X, \theta)\, p(\theta \mid D)\, d\theta$

Campionando dalla posteriore dei parametri e propagando:
$E[\lambda_m^X \mid D] \approx 11{,}47°$ con SD $\approx 0{,}52° \rightarrow$ IC_95% = [10,45°, 12,49°]
$E[\varphi_m^X \mid D] \approx 43{,}62°$ con SD $\approx 0{,}31° \rightarrow$ IC_95% = [43,01°, 44,23°]
Queste coordinate corrispondono approssimativamente alla regione tra La Spezia e Lucca nell'Italia nord-occidentale. Se esistesse evidenza archeologica o testuale di una città antica in quella zona con coordinate tolemaiche 35°, 43°, potremmo usare questa analisi per proporre identificazioni candidate.
Conclusioni Finali e Direzioni Future
Questa analisi esaustiva delle 27 città italiane nella Geografia di Tolomeo ha rivelato una struttura multi-scala degli errori cartografici:
Scala globale (~1000 km): Bias sistematico longitudinale dominante (~6° o ~500 km), compressione scala ~2,6×, catturato da trasformazione affine (R^2 ~ 0,82-0,91).
Scala regionale (~500 km): Bias specifici per isole (Sardegna +2,8° lat, Corsica +2,1° lat, Sicilia +1,2° lat), richiedono modellazione gerarchica, riducono RMSPE del ~10%.
Scala locale (~100 km): Distorsioni non lineari e idiosincrasie città-specifiche, richiedono splines o GP, spiegano ulteriore ~15% di varianza residua.
Scala micro (<50 km): Errori casuali di misura (~0,5° lat, ~0,8° lon), rappresentano limite fondamentale di accuratezza strumentale antica.
Le tecniche statistiche avanzate sviluppate - inferenza bayesiana MCMC, processi gaussiani multi-task, regressione robusta M-stimatori, modelli gerarchici - forniscono un template metodologico per analisi quantitative di cartografie storiche.
Questo lavoro dimostra che anche dati geografici bimillenari, affetti da errori sistematici massivi e strutturati in modi complessi, possono essere analizzati rigorosamente con strumenti statistici moderni, estraendo informazione quantitativa robusta sulla conoscenza geografica antica e sui processi di compilazione cartografica. La Geografia di Tolomeo emerge non come un semplice catalogo di errori, ma come un documento scientifico sofisticato che, pur con tutti i suoi limiti, rappresenta il culmine della cartografia matematica antica e un monumento alla capacità umana di costruire modelli quantitativi sistematici del mondo anche in assenza degli strumenti che oggi consideriamo indispensabili.